DOCUMENT RESUME

ED 468 955 TM 034 442

AUTHOR	Wilson, Kenneth M.; Powers, Donald E.
TITLE	Factors in Performance on the Law School Admission Test. Statistical Report. LSAC Research Report Series.
INSTITUTION	Law School Admission Council, Newtown, PA.
REPORT NO	LSAC-SR-93-04
PUB DATE	1994-03-00
NOTE	105p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC05 Plus Postage.
DESCRIPTORS	Ability; College Entrance Examinations; Higher Education; Law Schools; Law Students; *Performance Factors; Scores; *Test Construction; *Test Results
IDENTIFIERS	Graduate Record Examinations; *Law School Admission Test

ABSTRACT

This study was undertaken to clarify the internal structure of the Law School Admission Test (LSAT) and shed light on the ability or abilities measured by the three item types that make up the test (logical reasoning, analytical reasoning, and reading comprehension). The study used data for two forms of the LSAT for general samples of LSAT examinees and a subsample of LSAT examinees identified as also having taken the Graduate Record Examination (GRE) General Test, which includes the three item types that make up the LSAT. Within the joint LSAT/GRE sample, it was possible to conduct separate, parallel factor analysis of the parcels of item types common to both tests, assess time-related effects on between-test correlations, and analyze factor structure underlying performance on combined LSAT and GRE parcels of logical reasoning, analytical reasoning, and reading comprehension items. Unique features of the study were the use of data from both the LSAT and GRE General Test to assess structure in item types common to both tests and the method of pooling different items of the same type and position across multiple test forms in order to create parcels used to generate correlations for analysis. Study findings suggest a similar structure for logical reasoning, analytical reasoning, and reading comprehension item types regardless of the test in which they are used. The structure entails two dimensions, one involving aspects of general reasoning as measured by logical reasoning and reading comprehension item types, and the other more narrowly constrained, formal-deductive aspects of reasoning tapped by the analytical reasoning item type. (Contains 9 exhibits, 4 figures, 28 tables, and 35 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

LSAC RESEARCH REPORT SERIES

TM

ED 468 955

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

J. VASELECK

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Factors in Performance on the Law School Admission Test

Kenneth M. Wilson
Donald E. Powers

Law School Admission Council
Statistical Report 93-04
March 1994

TM034442

LAW Services

A Publication of Law School Admission Council/Law School Admission Services

The Law School Admission Council is a nonprofit association of United States and Canadian law schools. Law School Admission Services administers the Council's programs and provides services to the legal education community.

LSAT® and the Law Services logo are registered by Law School Admission Services, Inc. Law School Forum is a service mark of Law School Admission Services, Inc. *The Official LSAT PrepTest*, *The Official LSAT PrepKit*, and *The Official LSAT TriplePrep* are trademarks of Law School Admission Services, Inc.

Copyright© 1994 by Law School Admission Services, Inc.

All rights reserved. This book may not be reproduced or transmitted, in whole or in part, by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission of the publisher. For information, write: Publications, Law School Admission Services, Box 40, 661 Penn Street, Newtown, PA 18940.

This study is published and distributed by the Law School Admission Council (LSAC) and Law School Admission Services (LSAS). The opinions and conclusions contained in this report are those of the author and do not necessarily reflect the position or policy of LSAC/LSAS.

EXECUTIVE SUMMARY

Typically, a test—such as the Law School Admission Test (LSAT), the primary focus of this study—employs more than one type of item and/or different kinds of content for a given item type. For these multi-faceted measures, the extent to which the different types of test questions tap different aspects of a particular ability, or in fact tap several different abilities, needs to be assessed. In the case of the LSAT, as presently constituted, this issue does not appear to have been resolved. As described in Section 1 of the main report, this study was undertaken to clarify the internal structure of the LSAT, and shed light on the nature of the ability or abilities measured by the three types of test items that make up the LSAT—that is, reading comprehension, logical reasoning, and analytical reasoning.

The study drew on data for two different forms of the LSAT, namely, the June 1991 form and the October 1991 form. For broader perspective, the study also used data from the same two LSAT administrations, for a subsample composed of LSAT examinees identified through file-matching procedures as having taken the Graduate Record Examinations (GRE) General Test, between October 1988 and December 1991, inclusive. Time interval between GRE and LSAT testing occasions, without regard to order, ranged from five days to 36 months.

Items of the same types as those used in the current version of the LSAT have been included in all editions of the GRE General Test since October 1981. Thus, it was possible to draw on the substantial body of evidence generated in the GRE context regarding relationships among these item types. This research is reviewed in Section 2 of the report. Section 2 also includes information pointing up (a) strong similarities in the "surface characteristics" of the three item types as reflected in descriptions of the three item types in testing program publications, and illustrative items from the LSAT and the GRE, as well as (b) differences between the LSAT and the GRE with respect to internal organization of test items, number of sections, and so on.

Within the joint LSAT/GRE sample it was possible to

(a) conduct parallel **within-test** analyses of correlations among the three item types,

(b) assess time-related attenuating effects in patterns and levels of **between-test** correlations involving scaled scores and specially computed item-type subscores for the three common item types, and ultimately, by using combined data from both the LSAT and the GRE,

(c) assess the extent to which patterns of correlations involving parcels of items of three types common to both tests in the **combined LSAT/GRE sample** were similar to those identified in the two separate within-test analyses.

BEST COPY AVAILABLE

A series of related analyses, was undertaken.

First, separate analyses were made of within-test correlations involving primarily scores on item-type parcels (sets of four to six items of the same type) using

(a) LSAT data for a general sample of LSAT examinees and for the selected sample, alluded to above, of LSAT examinees who also took the GRE General Test (see Section 3), and

(b) GRE data for the selected LSAT/GRE sample, as reported in Section 4.

A unique feature of the study was the method of pooling different items of the same type and position across multiple forms to create parcels used to generate correlations for analysis.

Findings were generally parallel for the separate exploratory within-test factor analyses based on parcels of items of the three types common to both tests. In each analysis, when two factors were extracted, LSAT logical reasoning and reading comprehension items defined one factor, while LSAT analytical reasoning parcels defined the other. In the case of the LSAT, findings for the selected sample who took the GRE were in all essential respects similar to those for the general sample.

These parallel within-test findings suggested in both tests these item types measure psychometrically distinguishable aspects of reasoning ability: aspects of general or informal reasoning, defined by reading comprehension items and logical reasoning items, on the one hand, and aspects of formal, deductive reasoning, defined by analytical reasoning items, on the other.

Next, analyses were made of between-test correlations involving reported, scaled scores and specially computed item-type section scores. These analyses (described in detail in Section 5) were designed in part to assess effects associated with the fact that the LSAT observations and the GRE observations were collected on different testing occasions separated by intervals ranging from less than 10 days to 36 months. The between-test analyses included assessment of time-related effects on between-test correlations involving the three item types common to both tests.

In these analyses, profiles of correlations involving LSAT item types and their GRE counterparts, computed for shorter-interval (between tests) and longer-interval subgroups (< 10 days versus 19–36 months) were found to be strikingly similar with respect to pattern. They differed only with respect to level.

Results of the analysis of between-test correlations—observed and corrected for attenuation due to the presence of measurement error—involving item types common to both tests were consistent with the findings of the separate within-test factor analyses. In both instances findings suggested psychometrically distinguishable differences between aspects of general or informal reasoning measured by reading comprehension and logical reasoning item types, on the one hand, and aspects of formal, deductive reasoning tapped by the analytical reasoning items, on the other.

Last in the series of analyses, as described in Section 6, intercorrelations involving the combined set of LSAT and GRE logical reasoning, reading comprehension, and analytical reasoning item-type parcels were analyzed. The sets of LSAT and GRE parcels used in the two separate analyses were combined to produce intercorrelations for these analyses. The eigenvalues suggested that two factors, one defined primarily by LSAT and GRE logical reasoning and reading comprehension parcels and the other by LSAT and GRE analytical reasoning parcels, sufficiently characterized the correlational structure for the combined parcels.

Some Implications

In essence, the study findings suggest a common underlying structure for logical reasoning, reading comprehension, and analytical reasoning item type regardless of the test (LSAT or GRE) in which they are used. The structure appears to involve two dimensions. One dimension is represented by logical reasoning and reading comprehension item types measuring general reasoning skills that appear to be associated with the analysis of extended discourse. The other dimension represents a more narrowly constrained, formal-deductive aspect of reasoning, measured by the analytical reasoning item type.

For the LSAT, which currently reports only a single score to summarize performance involving three different item types, perhaps the most central conclusion supported by the findings is that

- the logical reasoning, reading comprehension, and analytical reasoning item types included in the LSAT, have potential to generate more information than is now being conveyed by the single LSAT scaled score.

That potential is suggested by the finding that the LSAT item types measure psychometrically distinguishable aspects of reasoning ability. This raises the attendant possibility that the information provided by item-type subscores might prove to be useful for predictive or diagnostic purposes in the LSAT context.

Questions concerning differential and/or incremental validity of subscores that might be computed are of immediate interest. For example, one score based on logical reasoning and reading comprehension items and a second score based on the analytical reasoning items would be consistent with the basic two-factor outcomes. Would the use of two scores, or perhaps a score for each LSAT item type, result in improved prediction of first-year law school grades generally, grades in particular courses or clusters of courses, grades in second-year courses?

A study of the comparative validity of LSAT subscores such as those noted above, for predicting such criteria—in general samples, and in samples defined by ethnic group membership, age, gender, undergraduate major, and so on—would contribute toward resolution of academically, psychometrically, and socially important "differential validity" questions in the current LSAT context.

Closely related to the foregoing are questions concerning incremental validity. For example, does differential weighting of item-type subscores result in better prediction of pertinent criteria (for example, grades in successive years of legal education), than is provided by a total scaled score based on a simple sum of section scores, in general samples of law students? In subgroups defined by ethnicity, gender, age, undergraduate major, and so on?

Study findings indicate that between-test correlations involving LSAT reading comprehension, logical reasoning, and analytical reasoning subscores and corresponding GRE subscores are differentially resistant to time-related attenuating influences. These findings suggest the possibility of differences in relative stability for the abilities involved. In the present study, inferences about relative stability, of course, are based on between-test correlations for the respective item-type subscores. It seems important to make a direct assessment of test-retest stability, short- and long-term, for LSAT item-type subscores.

These and other study findings have incidental implications for the GRE context. For example, the findings tend to confirm and extend conclusions based on GRE studies, namely, that logical reasoning items and analytical reasoning items are measuring psychometrically distinguishable aspects of reasoning ability. Thus, research questions such as those raised above for the LSAT, also have implications for continued research in the GRE context.

That LSAT logical reasoning, reading comprehension, and analytical reasoning items and their GRE counterparts have a common factor structure is important because this finding suggests that future research involving these item types in the LSAT context can draw on relatively extensive GRE research findings (such as those summarized briefly in Section 2 of the main report) both for formulating working hypotheses and for evaluating LSAT research outcomes. It also follows that as LSAT research findings involving these item types accrue, the LSAT findings in turn can usefully inform research in the GRE context.

Jointly planned research projects involving item types common to both tests might expedite attainment of objectives common to both testing programs—for example, clarifying distinctions between logical reasoning and reading comprehension.

In this connection, given the observed affinity between logical reasoning and reading comprehension—combined with hints of distinctiveness—it is noteworthy that the version of the logical reasoning item type considered in this study for both the LSAT and the GRE involves heavy reading comprehension requirements.

Accordingly, logical reasoning and reading comprehension are "linked" to some degree by heavy reading demands. To the extent that it is possible to measure "logical reasoning" using item types with limited reading demands, progress may be made in clarifying distinctions between "logical reasoning" and "reading comprehension," by cooperative research projects involving experimental logical reasoning and analytical reasoning items, and operational items from both tests, along lines followed in GRE research conducted by Emmerich, Enright, Rock, and Tucker (1991), for example.

In reaching decisions regarding score definition and score reporting, the results of this study involving data from both tests, suggest that both testing programs might benefit from research projects capitalizing on the common structure that appears to underly performance on the three types of items that are common to both the LSAT and the GRE General Test.

ABSTRACT

This study was undertaken to clarify the internal structure of the Law School Admission Test (LSAT), and shed light on the ability or abilities measured by the three item types that make up the test (logical reasoning, analytical reasoning, and reading comprehension). The study used data for two forms of the LSAT in (a) general samples of LSAT examinees, and (b) a subsample of LSAT examinees identified as having also taken the GRE General Test (GRE), which includes the three item types that make up the LSAT.

Within the joint LSAT/GRE sample, it was possible to conduct separate, parallel factor analyses of parcels of item types common to both tests, assess time-related effects on between-test correlations, and analyze factor structure underlying performance on combined LSAT and GRE parcels of logical reasoning, analytical reasoning, and reading comprehension items. Unique features of the study were (a) the use of data from both the LSAT and the GRE General Test to assess structure in item types common to both tests, and (b) the method of pooling different items of the same type and position across multiple test forms in order to create parcels used to generate correlations for analysis.

Study findings suggested a similar structure for logical reasoning, analytical reasoning, and reading comprehension item types, regardless of the test (LSAT or GRE) in which they are used. The structure entails two dimensions—one involving aspects of general reasoning as measured by logical reasoning and reading comprehension item types, and the other more narrowly constrained, formal-deductive aspects of reasoning tapped by the analytical reasoning item type.

Implications of the findings are discussed.

ACKNOWLEDGMENTS

This project was sponsored by the Law School Admisson Council whose support is acknowledged with appreciation. Lynne Norton, Law School Admission Services, Inc. (LSAS), developed the special LSAT files that were used in the study, and Kathleen McGeady, also a member of the LSAS staff, provided administrative assistance and guidance throughout.

At ETS, several colleagues contributed to the project. The critical task of matching LSAT files with GRE files in order to identify LSAT examinees who also had taken the GRE was completed by Nancy Robertson. James Ferris provided "common structure" for a file containing differentially ordered sets of GRE items from 37 different test administrations involving 21 generic forms of the GRE General Test. Ned Walthall and Manfred Steffen provided internally generated LSAT-related materials; Elana Broch, Raymond Thompson, and Karen Wisnia provided a variety of GRE-related materials. Very helpful reviews of a draft of this report were provided by Neil Dorans and Spencer Swinton.

These contributions are acknowledged with appreciation. However, the writers assume full responsibility for the contents of the report.

Table of Contents

EXECUTIVE SUMMARY i

ABSTRACT vi

ACKNOWLEDGMENTS vii

SECTION 1. INTRODUCTION 1
    Purpose of the Present Study 2
    Overview 3
        Analyses of the Internal Structure of the LSAT 3
        Assessing GRE Dimensionality in the Selected LSAT/GRE Sample 3
        Evaluation of Between-Test Correlations 4
        Exploratory Factor Analyses Involving Parcels of Item Types Common to Both Tests 5

SECTION 2. REVIEW OF RELATED RESEARCH IN THE GRE CONTEXT 6
    Characteristics of Item Types Common to Both Tests 11

SECTION 3. ASSESSING THE DIMENSIONALITY OF THE LSAT 21
    General Analytical Rationale 22
    Study Sample and Data 24
    Analytical Approach 27
    Analytic Procedure and Related Findings 28

SECTION 4. ANALYSIS OF FACTORS UNDERLYING PERFORMANCE ON THE GRE GENERAL TEST IN THE SELECTED LSAT/GRE SAMPLE 34
    The GRE Data Set 34
        Aggregating Data Across GRE Forms 35
    Analytic Procedure 39
    LSAT-Parallel Analyses 50

SECTION 5. BETWEEN-TEST CORRELATIONS IN THE JOINT SAMPLE, AND CONSIDERATIONS INVOLVED IN THEIR EVALUATION 52
    Scaled Score Analyses 54
    Patterns of Between-Test Correlations Involving Subscores on Item Types Common to Both Tests 56
        Item-Type Subscores Involved in Analyses 57
    Time-Related Trends in Observed LSAT/GRE Correlations 60
    Estimating Between-Test Overlap 68

Table of Contents continued

SECTION 6. FURTHER EVIDENCE OF COMMON STRUCTURE IN ITEM TYPES COMMON TO THE LSAT AND THE GRE 71
Analytical Rationale and Procedure 73
Findings 75
Evidence of Common Structure 78

SECTION 7. STUDY FINDINGS AND IMPLICATIONS: REVIEW AND EVALUATION 80
Recapitulation 81
Implications 83

REFERENCES 87

Tables

Table 2.1 Summary of Factor Analytic Studies of the GRE General Test . . . 7

Table 2.2 General Organization of Tasks in the LSAT and in the GRE General Test, Respectively . . . 20

Table 3.1 Organization of the LSAT . . . 21

Table 3.2 Observed Correlations and Correlations After Correction for Attenuation: LSAT Section Scores, June 1991 and October 1991 Forms of the LSAT . . . 23

Table 3.3 Descriptive Statistics for General and Selected LSAT Samples, by LSAT Form Taken: Designated Variables . . . 26

Table 3.4 General Definitional Characteristics of Parcels of LSAT Items . . . 27

Table 3.5. Summary of Principal Components Extractions in Six LSAT Analysis-Samples: Three Factors . . . 30

Table 3.6 Varimax Factors for Two-Factor and Three-Factor Solutions in the General LSAT, Combined-Forms Analysis Sample (N = 7,385) . . . 31

Table 3.7 Pattern Coefficients for LSAT Parcels on Oblimin Factors: Two-Factor and Three-Factor Solutions in the General LSAT, Combined-Forms Sample (N = 7,385) . . . 32

Table 3.8 Summary of Factor Correlations for Two-Factor and Three-Factor Oblimin Solutions in the Six Analysis Samples . . . 33

Table 4.1 Distribution of GRE Forms, by LSAT Form Taken . . . 36

Table 4.2 Distribution of the LSAT/GRE Sample by GRE Testing Date . . . 36

Table 4.3 Description of GRE Item-Type Parcels Used in the Analysis . . . 38

Table 4.4 Summary of Eigenvalues and Associated Percentages of Total Variance . . . 40

Table 4.5 Loadings of GRE Parcels on Three Orthogonal Factors, and Corresponding Pattern Coefficients on Three Oblique Factors: Three-Factor Model . . . 41

Table 4.6 Loadings of GRE Parcels on Factors: Four-Factor Solutions . . . 43

Tables continued

Table 4.7 Factor Correlations: Emmerich et al. and Present Study . . . 49

Table 4.8 Eigenvalues and Associated Percentages of Total Variance: Three Principal Components for GRE Reading Comprehension, Logical Reasoning, and Analytical Reasoning Parcels . . . 50

Table 4.9 Two- and Three-Factor Orthogonal (Varimax) and Oblique (Direct Oblimin) Solutions: GRE Reading Comprehension, Logical Reasoning, and Analytical Reasoning Parcels . . . 51

Table 5.1 Distribution of Sample According to Time Between LSAT and GRE Test Administrations and Ordering of the Two Test Administrations . . . 53

Table 5.2 Descriptive Statistics for Test-Order/Time-Interval Subgroups: LSAT and GRE Scaled Scores . . . 55

Table 5.3 Correlation of GRE General Test Scores with LSAT Score, By Order Time/Interval Categories: Scaled Scores for Both Tests . . . 56

Table 5.4 Outline of Item-Type Subscores from the LSAT and the GRE . . . 58

Table 5.5a Estimates of Reliability for LSAT Subscores . . . 59

Table 5.5b Estimates of Reliability for GRE Item-Type Part Scores, Using Multiple-Form Data and Single-Form Data, Respectively . . . 60

Table 5.6 Correlation of LSAT Scaled Score and Item-Type Section Scores with Designated GRE Scores, in Nine Order/Interval Subgroups . . . 62

Table 5.7 Summary of Observed and Disattenuated Correlations between Selected LSAT Variables and Designated GRE Variables, for Shorter- and Longer-Interval Subgroups . . . 69

Table 6.1 Summary of Selected Results of Principal Components Analysis: Eigenvalues and Associated Percentage of Total Variance: Six Components . . . 74

Table 6.2 Pattern Coefficients for LSAT and GRE Parcels on Oblique (Correlated) Factors in Separate Analyses and in Analyses Involving Combined Parcels from Both Tests . . . 79

Figures

Figure 5.1 Correlation of LSAT reading comprehension, logical reasoning, and analytical reasoning subscores with GRE item-type subscores for subgroups classified by time between tests: < 10 days vs. 19–36 months 64

Figure 5.2 Correlation of LSAT logical reasoning (LLR), reading comprehension (LRC), and analytical reasoning (LAR) subscores with GRE item-type subscores 66

Figure 5.3 Correlation of LSAT item-type subscores and corresponding GRE item-type subscores with designated GRE verbal and quantitative item-type subscores 68

Figure 5.4 Corrected correlations between LSAT and GRE subscores for item types common to both tests 70

Exhibits

Exhibit A1. Description of the Logical Reasoning Item Type: LSAT and GRE Versions 12

Exhibit A2. Illustrative Logical Reasoning Items: LSAT and GRE Versions 13

Exhibit B1. Description of the Analytical Reasoning Item Type: GRE and LSAT Versions 14

Exhibit B2. Illustrative Analytical Reasoning Items: LSAT and GRE Versions 15

Exhibit C1. Description of the Reading Comprehension Item Type: LSAT and GRE Versions 16

Exhibit C2a. Illustrative GRE Reading Comprehension Items (Abridged Options) 17

Exhibit C2b. Illustrative LSAT Reading Comprehension Passage and Questions (Abridged Options) 18

Exhibit 4a. Selected Findings of a Study by Emmerich, Enright, Rock, and Tucker (1991): Table 17a from the Study Report 46

Exhibit 6a. General Pattern of Factorial Decomposition of Intercorrelations of LSAT and GRE Parcels as Indicated by Item Types Defining Factors Identified in Two-, Three-, Four-, Five-, and Six-Factor Oblique Solutions 76

FACTORS IN PERFORMANCE ON THE LAW SCHOOL ADMISSION TEST

SECTION 1. INTRODUCTION

The Law School Admission Test (LSAT) is described (Law School Admission Services, 1991, p. 2) as being designed

> ". . . to measure skills that are essential for success in law school: the ability to read and comprehend complex texts with accuracy and insight, organize and manage information and draw reasonable inferences from it, reason critically, and analyze and evaluate the reasoning and argument of others. These skills are typically developed over a long period of time. The LSAT provides candidates and law schools with a useful measure of a candidate's ability to handle new information, using well-developed reading skills and careful and logical thought as the principal tools."

The skills alluded to above are assessed operationally by three different types of multiple-choice test questions or items, labelled, **reading comprehension, analytical reasoning,** and **logical reasoning**, described (for example, LSAT, 1991), as follows:

- Reading comprehension questions require test takers to read carefully and accurately, to determine the relationships among the various parts of the passage, and to draw reasonable inferences from the material in the passage.

- Each logical reasoning question requires the examinee to read and comprehend the argument or the reasoning contained in a short passage, and then answer one or two questions about it. The questions test a variety of abilities involved in reasoning logically and critically (including, for example,) drawing reasonable conclusions from given evidence or premises.

- Analytical reasoning items are designed to measure the ability to understand a structure of relationships and to draw conclusions about the structure. The examinee is asked to make deductions from a set of statements, rules, or conditions that describe relationships among entities such as persons, places, things, or events.

These three item types have been used in all forms of the LSAT developed since 1982, and are treated operationally as measuring related aspects of one general underlying ability. For example, test performance is summarized by a single, LSAT scaled score based on the total number of correct responses to test items, without regard to type.

Purpose of the Present Study

If a test employs more than one type of item and/or different kinds of content for a given item type, the extent to which the different types of test questions tap different aspects of a particular ability, or in fact tap several different abilities, needs to be assessed. In the case of the LSAT as presently constituted, this issue does not appear to have been resolved.[1]

The present study was undertaken to clarify the internal structure of the LSAT, and shed light on the ability or abilities measured by LSAT reading comprehension, logical reasoning, and analytical reasoning item types. The study drew on data from two regularly scheduled LSAT administrations, namely, the June 1991, and October 1991 administrations.

For broader perspective, the study also draws on data for a subsample from the same two LSAT administrations, composed of LSAT examinees identified **as having also taken the GRE General Test** (for example, ETS, 1990) between October 1988 and December 1991. Using data for this selected sample it was possible to conduct parallel within-test analyses involving item types common to both tests, examine between-test correlations involving scaled scores and specially computed item-type subscores, and ultimately to identify factors underlying performance on parcels of items of three types that are common to both tests, using combined LSAT and GRE data.

The study also drew on the substantial body of evidence regarding relationships among these item types based on research in the GRE context (to be reviewed later) involving GRE reading comprehension, logical reasoning, and analytical reasoning items similar to those used in the LSAT. These three item types have been included in all editions of the GRE since October 1981.

Detailed descriptions of study data and the analytical procedures used to attain study objectives, as well as related findings, are provided in the remaining sections of this report. A general overview of the organization of the report and the general analytical approach employed in the study, is provided below.

---

[1] A review of Volumes I-IV of Reports of LSAC Sponsored Research (LSAC, 1976a, 1976b, 1977, 1984) reveals only one study (Carlson, 1976b)--authorized by the LSAC in 1966 and reported in 1970--concerned with identifying the factor structure of an extant version of the LSAT in order to answer the question: What are the abilities being measured by LSAT items? According to Carlson (p. 12), "It was hoped that a classification of items based on abilities would provide a useful supplement or alternative to the present classification based on item type . . . ." Carlson factored matrices of interitem coefficients for items of the eight different types included in the 1966 version of the LSAT, and identified seven factors. The first factor, called "verbal ability," was defined primarily by reading comprehension items; the second factor, called "verbal inductive reasoning," was defined primarily by a subset of "Principles and Cases" items." The "Principles and Cases" item type used until June 1982, was ". . . developed to test the candidate's ability to reason logically" (McPeek, Pitcher, and Carlson, 1976b, emphasis added). During the course of the present study, another LSAC-sponsored factor analytic study (Camilli, Wang, and Fesq, 1992) involving several forms of the LSAT used between June 1989 and October 1990--the same version of the LSAT as that under consideration herein--was reported. Findings reported herein for LSAT forms used in the June 1991 and October 1991 administrations are consistent with those reported by Camilli, Wang, and Fesq (1992) for the forms used earlier.

Overview

Section 2 provides (a) a detailed review of previous research findings regarding the internal structure of the GRE General Test, and the nature of its constituent item types, especially GRE logical reasoning, reading comprehension, and analytical reasoning, and (b) a brief comparison of the composition of both the LSAT and the GRE.

Analyses of the Internal Structure of the LSAT

Section 3 describes analyses of the internal structure of the LSAT—that is, its dimensionality, the principal concern of the present study. Dimensionality was assessed by analyzing intercorrelations of subscores based on LSAT reading comprehension, logical reasoning, and analytical reasoning items, and smaller subsets (parcels) of LSAT items within each item type. For example, exploratory factor analyses were conducted using data for (a) general LSAT samples taking different forms of the LSAT used in the June 1991 and October 1991 test administrations, and (b) the selected subsample from those two test administrations for whom GRE data were also available.

Analyses involving LSAT performance for the selected subsample of LSAT examinees (those who also took the GRE) yielded results that were consistent with the outcomes observed for the general LSAT samples. The similarity in outcomes suggested, among other things, that data for the selected LSAT sample also provide a sound basis for drawing inferences about LSAT dimensionality.

Assessing GRE Dimensionality in the Selected LSAT/GRE Sample

Section 4 describes (a) an analysis of factors underlying performance on a full array of GRE item-type parcels, designed in part to shed light on the extent to which the GRE data adequately represent the GRE test taking population, and (b) an analysis of factors underlying performance on parcels of items of the three types common to both the LSAT and the GRE.

The possibility of unrepresentative findings due to selected samples is a pertinent consideration generally. Questions regarding representativeness are especially germane here because the members of the selected LSAT/GRE sample took some 21 different generic forms of the GRE, on over 30 different testing occasions spanning a total of 42 months.

GRE data for a general (unselected) sample of GRE examinees were not included in the study. Thus it was not possible to assess directly the issue of representativeness of findings based on the selected GRE data set, as was done in the case of the LSAT data. Accordingly, the issue was addressed indirectly by evaluating the results of exploratory analyses designed to identify factors underlying performance on GRE items of all nine types, in the selected sample, in light of previous research. These analyses were based on pooled within-form correlations in data aggregated across 21 generic forms of the GRE.

Results of the exploratory analyses, involving a full array of GRE item-type parcels, conformed generally to previously reported patterns, especially that identified by Emmerich, Enright, Rock, and Tucker (1991) in research involving both operational and experimental logical reasoning and analytical reasoning GRE item types. The orderly nature of GRE findings in the selected sample indicated that data for this sample provided a sound basis for drawing inferences regarding GRE dimensionality generally, as well as the dimensionality of the three item types that are common to both the LSAT and the GRE.

Based on the foregoing, analyses of factor structure involving parcels of GRE reading comprehension, logical reasoning, and analytical reasoning items were conducted. These analyses parallel those involving the three LSAT counterparts, alluded to above, and the findings of the two within-test analyses also were generally parallel.

Section 4 also provides information regarding the characteristics of the joint LSAT/GRE sample, the nature of the GRE data involved and the procedures that were used in aggregating data across forms and testing occasions, as well as the factor analyses outlined above.

Evaluation of Between-Test Correlations

The correlational findings reported in Section 3, involving LSAT data only, and Section 4, involving GRE data only, reflect concurrent, within-test relationships involving parcels of test items—including, in the case of the GRE, sets of concurrent, within-test observations based on different forms of the test, taken on 37 separate testing occasions.

Among other things, these findings suggested that outcomes of correlational analyses involving GRE data only, as well as those involving LSAT data only, were not unduly influenced by (potentially interesting and probably complexly interacting) variables associated with selection into the LSAT/GRE sample.

However, as described in detail in Section 5, in all instances, the LSAT and the GRE were taken on separate testing occasions under different testing conditions. The time interval between the two tests ranged from less than 10 days to approximately 36 months, and the order of test taking was not the same for all members of the joint sample.

Thus, correlations between LSAT and GRE scores (total scores, item scores, parcel scores, and so on) may be affected to some extent by factors such as the following:

(a) conditions affecting performance on one testing occasion but not the other (anxiety, luck, illness, preparation or lack of preparation, shifts in motivation, and so on),

(b) differential change (growth or decline) in the abilities being tapped by both tests, due to individual differences in experiences (over periods of up to three years between tests),

(c) "short-term practice effects," due to proximity of test dates; and so on.

Analyses of between-test correlations reported in Section 5, revealed time-related trends in correlations involving both scaled total scores and specially computed subscores for reading comprehension, logical reasoning, and analytical reasoning items. For example, observed between-test relationships were stronger in data for subgroups taking the two tests within 10 days, than in data for subgroups in which time between testing occasions was 19-36 months. In addition, when classified according to time interval between test occasions and sequence of testing, subgroups differed significantly with respect to average test performance—true for both the LSAT and the GRE.

The between-test correlational findings involving item-type subscores from the LSAT and the GRE suggested the possibility of differences in relative stability over time of the abilities tapped by the three item types—for example, the possibility of differential change in the abilities being measured. Explicating the observed time-related trends and/or evaluating the relative influence of particular variables that may have affected observed relationships between LSAT and GRE scores is outside the scope of the present study.

An evaluation is made of both observed correlations between LSAT variables (scaled scores, item-type section scores, and so on) and GRE variables, and correlations after correction for attentuation due to measurement error.

Exploratory Factor Analyses Involving Parcels of Item Types Common to Both Tests

Apart from their substantive implications, the findings reported in Section 5 indicate clearly that the interpretation of between-test relationships is complicated by the fact that performance on the two sets of test items was observed at different times and under different conditions.

Factor analyses involving the combined parcels of items of the three types that are common to both the LSAT and the GRE—that is, the parcels that were used in the separate within-test analyses—are reported in Section 6. Despite the interpretive complications alluded to above, the analyses shed additional light regarding the classification, according to ability domain, of reading comprehension, logical reasoning, and analytical reasoning item types for both the LSAT and the GRE.

Findings of the analysis of combined parcels of the three item types were generally similar to the findings of parallel within-test analyses. On balance, these findings suggest that a common factor structure underlys performance on the three types of items that are common to the LSAT and the GRE, regardless of the test for which they are developed.

Section 7 summarizes study findings and suggests some implications for both the LSAT and the GRE contexts.

SECTION 2. REVIEW OF RELATED RESEARCH IN THE GRE CONTEXT

As indicated in Section 1, reading comprehension, logical reasoning, and analytical reasoning items, similar in type to the corresponding LSAT item types, are used in forms of the Graduate Record Examinations (GRE) General Test. In the GRE context, these item types are treated operationally as belonging to two different ability domains, namely, an analytical ability domain and a verbal ability domain, respectively.

For the GRE, "analytical ability" is defined operationally by the logical reasoning and analytical reasoning item types, which are characterized as measuring somewhat different aspects of ". . . the ability to think analytically" (ETS, 1990, p. 41). "Verbal ability" is defined by the reading comprehension item type and three other verbal item types—analogy, antonym and sentence completion questions—characterized collectively as measuring aspects of ". . . ability to reason with words in solving problems . . . (that is) the ability to discern, comprehend, and analyze relationships among words or groups of words and within larger units of discourse such as sentences and written passages (and so on)" (ETS, 1990, p. 31).

Separately scaled number-right scores are reported for GRE analytical ability and GRE verbal ability, respectively, as well as for GRE quantitative ability.[2]

GRE "quantitative ability" is defined by quantitative comparison, regular mathematics, and data interpretation item types which are characterized as measuring . . . basic mathematical skills, understanding of elementary mathematical concepts, and ability to reason quantitatively and to solve problems in a quantitative setting" (ETS, 1990, p. 36).

Results of factor analyses involving GRE verbal, quantitative, and analytical items or item parcels, summarized in Table 2.1, generally tend to support a three-factor hypothesis, corresponding to divisions of the test for which separate scores are reported. That is, the items appear to measure psychometrically distinguishable verbal, quantitative, and analytical abilities—more clearly so for items in the GRE verbal domain and the GRE quantitative domain, than for items in the GRE analytical ability domain.

---

[2] In contrast to legal education, in which the first-year curriculum tends to be relatively uniform from school to school, graduate education is "notoriously" diverse. This fundamental difference is reflected in the different approaches to admissions testing for graduate education and legal education. The GRE General test, which reports three scores, must be appropriate for disciplines that differ on a number of grounds, e.g., the extent to which they are verbally or quantitatively oriented. The LSAT, not being constrained in this way, summarizes performance on three different item types in a single score in order to measure a more general reasoning ability.

Table 2.1. Summary of Factor Analytic Studies of the GRE General Test

Study	Sample	Test Version*	Methods	Major Results
Powers, Swinton, & Carlson (1977)	Two random samples each consisting of 8,000 examinees who took the two operational test forms given in October 1975.	Version used until October 1977.	Exploratory analysis of item-level data, analysis of tetrachoric correlation matrices, orthogonal rotation.	Three major uncorrelated factors were extracted from each of the two test forms analyzed: a general quantitative factor, a reading comprehension factor, and a vocabulary factor. Several other much less prominent dimensions were also noted.
Powers & Swinton (1981)	Two random samples--one of 4,000 and the other of 6,000--who took two forms of the restructured test in October 1977.	Version used from October 1977 until October 1981.	Exploratory analysis in two stages, the first using item-level analyses of tetrachoric correlations and the second using clusters (parcels) of related items. Both orthogonal and oblique rotations were employed.	Three major uncorrelated factors emerged: verbal reasoning/reading comprehension, quantitative reasoning, and vocabulary. An analytical factor emerged as one of several other less prominent factors. Using an oblique rotation, four prominent factors were extracted: reading comprehension, general quantitative, vocabulary, and analytical. The correlations among these factors ranged from .37 for vocabulary with quantitative to .66 for reading with analytical.
Rock, Bennett, & Jirele (1986)	Three groups of disabled students: two groups of visually impaired examinees (N=188 and N=151)--one of which took a large-type, extended time administration--and one group of physically disabled examinees (N=108).	Version used after October 1981.	Confirmatory factor analysis.	A three-factor model--verbal, quantitative, and analytical--provided the best fit for each of the three handicapped groups. However, analytical scores did not have quite the same meaning for two disabled groups as they did for nondisabled examinees.
Rock, Werts, & Grandy (1982)	3,500 social science majors taking the test in September 1978.	Version used from October 1977 until October 1981.	Confirmatory factor analysis based on odd and even halves for each item type.	Several alternative models were fit. A three-factor solution (verbal, quantitative, and analytical) indicated that the analytical factor correlated .92 with the quantitative factor and .77 with the verbal factor. The correlation between the verbal and quantitative factors was .64. A slightly better fit was obtained when reading comprehension was considered as a fourth factor.

Study	Sample	Test Version*	Methods	Major Results
Stricker & Rock (1987)	Three samples each of 1,000 examinees at different age levels who took the GRE General Test in December 1982.	Version used after October 1981.	Confirmatory factor analysis based on parcels of items of the same type, using LISREL VI (Joreskog & Sorbom, 1981) as the method of analysis.	A three-factor solution provided the most acceptable fit. Factors were termed verbal, quantitative, and analytical abilities. At the three different age levels the verbal and quantitative factors correlated .53 to .63, the verbal and analytical .76 to .84, and the quantitative and analytical .74 to .80.
Schaeffer & Kingston (1988)	Several samples, each consisting of 1,000 examinees, were selected from five undergraduate majors (education, engineering, English, psychology or mathematics). These examinees took one form of the test that was administered in October 1984, April 1985, and December 1985.	Version used after October 1981.	Full-information factor analysis using the program TEST-FACT, which generates IRT parameters to estimate interitem correlations, which are in turn used as the basis for a principal factors analysis.	Prominent verbal and quantitative factors were extracted for each of the undergraduate major samples. A third less prominent, but significant, analytical factor also emerged for each sample. Correlations between verbal and quantitative factors ranged from .48 to .68; correlations between verbal and analytical factors ranged from .46 to .67; correlations between quantitative and analytical factors ranged from .52 to .78.
Emmerich, Enright, Rock, & Tucker (1991)	Two samples of approximately 370 examinees each, taking batteries with both operational and experimental GRE items, tested in April 1989. The study was designed to evaluate a number of experimental item types that might serve to clarify the analytical ability measure.	Version used after October 1981.	Confirmatory and exploratory factor analysis.	In confirmatory analyses, parcels of scores on operational GRE item types were constrained to have nonzero loadings only on factors they were expected to identify. In these analyses, loadings for logical reasoning on the analytical factor were considerably lower than were loadings for analytical reasoning. In an exploratory analysis involving both operational and experimental items, logical reasoning parcels tended to have higher loadings on the verbal factor than on the analytical factor, defined primarily by operational analytical reasoning parcels.

*The test used until October 1977 reported verbal and quantitative scores based on the following verbal and quantitative item types: verbal analogies, antonyms, sentence completion, reading comprehension, regular math (arithmetic, algebra, and geometry), and data interpretation. In October 1977 the analytical measure, which contained analysis of explanations, logical diagrams, and analytical/logical reasoning items, was introduced. An additional item type, quantitative comparisons, was introduced into the quantitative measure, and some shorter reading passages were substituted for larger ones in the verbal measure. In October 1981 the analytical measure was revised by deleting analysis of explanations and logical diagrams, and substituting greater numbers of analytical reasoning and logical reasoning items.

In the GRE verbal domain, two closely related subdomains have been identified relatively consistently, namely,

- "reading comprehension," a factor typically defined by the GRE reading comprehension item type—calling for discourse-level analysis—alone, or in combination with the sentence-completion item type that involves sentence-level analysis, and

- "vocabulary," a factor typically defined by analogy and antonym items, and (less consistently) by sentence-completion items—item types that do not require discourse-level analysis.

Stricker and Rock (1985), commenting on the fact that these two verbal subdomains did not emerge in their study, suggested that the outcome might be due to differences across studies in sample selection, factor methodology, and so on.

The GRE quantitative ability domain has been consistently and clearly defined by the quantitative comparison and regular mathematics item types. Data interpretation items typically have somewhat lower loadings than either of the other two item types on quantitative factors identified in these analyses, but nonetheless tend to be more closely identified with the quantitative ability domain than with either of the other GRE ability domains (see also Kingston and Dorans, 1982).

However, research findings indicate that in the GRE analytical ability domain, the alliance between component item types is an uneasy one. For example, GRE analytical reasoning and GRE logical reasoning items tend to have less common variance than do item types within the GRE verbal and GRE quantitative ability domains. Furthermore, what they have in common tends to overlap substantially with item types represented in either the verbal or the quantitative domain (for example, Stricker and Rock (1985: p. 25).

Schaeffer and Kingston (1988) found that a GRE analytical ability factor was defined principally by the GRE analytical reasoning item type, and concluded from this and other findings, as follows (p. 9):

> "This study has suggested that an analytical dimension in the GRE General Test may be defined by the analytical reasoning items and not by the logical reasoning items. This finding raises doubt about the utility of including analytical and logical reasoning items in the same score" (emphasis added).

In an analysis of GRE factor structure in two samples (Emmerich, Enright, Rock, and Tucker, 1991), in which three sets of operational GRE item types were constrained to have nonzero loadings only on the factors which they were expected to define, analytical reasoning parcels had strong loadings on the "analytical ability" factor (loadings ranged between .83 and .87), whereas those for logical reasoning parcels were comparatively considerably weaker (ranging between .38 and .56 across analyses in two samples [Emmerich et al., p. 38]).

In a more general, unconstrained exploratory analysis involving both operational and experimental logical reasoning and analytical reasoning GRE item types, two of three operational logical reasoning parcels loaded more heavily on the "verbal" factor than on a factor labelled "formal-deductive reasoning"—a factor defined primarily by very strong loadings for three operational GRE analytical reasoning parcels.

Results of studies involving subscores based on GRE item types (for example, Wilson, 1985; 1986) also indicate that GRE logical reasoning items and GRE analytical reasoning items may tap somewhat different underlying abilities.

For example, correlations between GRE logical reasoning and GRE analytical reasoning subscores tended to be *lower* than correlations between GRE logical reasoning and GRE verbal item-type subscores, and also lower than those between GRE analytical reasoning subscores and subscores on GRE quantitative item types.[3]

For examinees classified by major area of study, patterns of relative performance on GRE analytical reasoning and logical reasoning were consistent with expectation based on major-area differences in relative performance on quantitative and verbal items—that is, majors in primarily verbal fields had relatively higher means on GRE reading comprehension and logical reasoning subscores than on subscores for quantitative item types, while majors in quantitatively demanding fields performed relatively better on GRE quantitative comparison items and analytical reasoning items than on either GRE reading comprehension or logical reasoning items.[4]

Also, an ethnic group known to be characterized by higher means on GRE *quantitative* ability than on GRE verbal ability, namely, Asian Americans, had higher relative standing on GRE analytical reasoning than either GRE logical reasoning or GRE verbal items, *without regard to field*.[5]

Thus, research findings from the GRE context indicate that the logical reasoning item type appears to tap abilities that are closely akin to those measured by the reading comprehension item type, hence may belong to the verbal ability domain, rather than to the "analytical ability" domain. And, as suggested by Schaeffer and Kingston (1988), the "analytical ability" domain itself may be defined more validly by the analytical reasoning item type alone than by a combination of logical reasoning and analytical reasoning item types.

---

[3] Wightman and Muller (1990b) reported intercorrelations of item-type section scores for LSAT analytical reasoning, logical reasoning, and reading comprehension from a recent version of the LSAT as follows: analytical reasoning and reading ($r = .49$), logical reasoning and reading ($r = .67$), and analytical reasoning and logical reasoning ($r = .53$).

[4] Pitcher (1977b, Table C.2, pp. 480-481) reported means for law students grouped by undergraduate major field, on experimental logical reasoning and quantitative comparison items, as well as on operational data interpretation and reading items. Profiles of logical reasoning and reading means were quite similar across fields, as were profiles of data interpretation and quantitative comparison means. In both cases, the profiles conformed to a pattern consistent with expectation based on differences among the fields involved with respect to verbal-relative-to-quantitative emphasis in curricular content.

[5] Means and standard deviations for LSAT section scores were reported by Wightman and Muller (1990a) for generally comparable ethnic subgroups. The major-field mix in the sample was not reported. However, Asian Americans registered the highest relative standing on AR (mean = 18.83, approximately .15 standard deviation units above the estimated grand mean of 17.795); z-scaled mean on LR was -.16.

Characteristics of Item Types Common to Both Tests

Published descriptions of the abilities tapped by LSAT reading comprehension, logical reasoning, and analytical reasoning items (for example, LSAT, 1991), are very similar to those published for the GRE counterparts (for example, ETS, 1990). Excerpts from these published descriptions, as well as illustrative logical reasoning, analytical reasoning, and reading comprehension items from the two testing contexts, are summarized in a series of exhibits (A1 through C2b). The exhibits are arranged to provide side-by-side descriptions of an item type from the two testing contexts, followed by a side-by-side display of examples of items of the type under consideration.

Exhibit A1. Descriptions of the Logical Reading Item Type: LSAT and GRE Versions

GRE General Test

LOGICAL REASONING

Logical reasoning questions test the ability to understand, analyze, and evaluate arguments. Some of the abilities tested by specific questions include

- recognizing the point of an argument,
- recognizing assumptions on which an argument is based,
- drawing conclusions and forming hypotheses,
- identifying methods of argument, evaluating arguments and counter arguments, and analyzing evidence.

Each question or group of questions is based on a short argument, generally an excerpt from the kind of material graduate students are likely to encounter in their academic and personal reading. Although arguments may be drawn from specific fields of study such as the humanities, social studies, and the physical sciences, materials from more familiar sources such as political speeches, advertisements, and informal discussions or dialogues also form the basis for some questions. No specialized knowledge of any particular field is required for answering the questions, however, and no knowledge of the terminology of formal logic is presupposed.

Specific questions asked about the arguments draw on information obtained by the process of critical and analytical reading described above.

Law School Admission Test

LOGICAL REASONING

Logical reasoning questions are designed to evaluate a test taker's aptitude for understanding, analyzing, critizing, and completing a variety of arguments. The stimulus passages in logical reasoning questions use material from a variety of sources, including letters to the editor, speeches, advertisements, newspaper articles and editorials, informal discussions and conversations, and articles in the humanities, the social sciences, and the natural sciences.

Each logical reasoning question requires the examinee to read and comprehend the argument or the reasoning contained in a short passage, and then answer one or two questions about it. The questions test a variety of abilities involved in reasoning logically and critically. These include:

- recognizing the point or issue of an argument or dispute;
- detecting the assumptions involved in an argument or chain of reasoning;
- drawing reasonable conclusions from given evidence or premises;
- identifying and applying principles;
- identitying the method or structure of an argument or chain of reasoning;
- detecting reasoning errors and misinterpretations;
- determining how additional evidence or argument affects an argument or conclusion; and
- identifying explanations and recognizing resolutions of conflicting facts or arguments.

The questions do not presuppose knowledge of the terminology of formal logic. For example, you will not be expected to know the meaning of specialized terminology of formal logic. For example, you will not be expected to know the meaning of specialized terms such as 'ad hominen' or 'syllogism.' On the other hand, you will be expected to understand and critique the reasoning contained in arguments. This requires that you possess, at a minimum, a college-level understanding of widely used concepts such as 'argument,' 'premise,' 'assumption,' and 'conclusion.'

Exhibit A2. Illustrative Logical Reasoning Items: LSAT and GRE Versions

GRE General Test

__Directions__. As for Analytical Reasoning (both types in same section) "Each question or set of quesions is based on a passage or set of conditions. In answering some of the questions, it may be useful to draw a rough diagram. For each question, select the best answer choice given."

7. The presence of microorganisms that produce a toxin causes seawater to turn brownish red, a phenomenon known as a red tide. Sea otters do not feed in areas where clams, their main source of food, have been contaminated with this toxin. According to a proposed explanation of the otters' behavior, the otters sample the clams in a potential feeding area and can taste any toxin in them.

Which of the following, if true, would most strongly indicate that the hypothesis described in the last sentence of the passage is __not__ correct?

(A) In some of the ares where red tides occur, neither clams nor sea otters are indigenous species.
(B) The presence of sea otters in a given area has a significant effect on which other marine organisms are to be found in that area.
(C) When seawater in an area unaffected by red tide is artificially dyed brownish red, sea otters do not feed on the clams in that area.
(D) If the clams in a given area are contaminated with toxins, sea otters move to other areas in search of food.
(E) Although very small amounts of the toxin produced during a red tide are not harmful, large doses can be fatal to animals the size of sea otters.

8. An acre of average farmland produces only about 400 pounds of grain amaranth, as against 2,400 pounds per acre, or six times as much, for wheat. It follows that whenever the grain-amaranth price is projected to be more than six times the projected price of wheat, farmers wishing to maximize profits will grow grain amaranth rather than wheat.

The argument above is based on which of the following assumption?

(A) An acre's worth of grain amaranth is no more expensive to grow and bring to market than an acre's worth of wheat.
(B) There is no crop that produces a higher yield in terms of pounds harvested per acre than wheat.
(C) By choosing which crops to grow, farmers can exert a significant influence on prices of those crops.
(D) Farmers are no less motivated by the desire to maximize profits than are other occupational groups.
(E) Prices of grain crops can change faster than farmers can change the acreage devoted to various grain crops.

Law School Admission Test

__Directions__. The questions in this section are based on the reasoning contained in brief statements or passages. For some questions, more than one of the choices could conceivably answer the question. However, you are to choose the __best__ answer; that is, the response that most accurately and completely answers the question. You should not make assumptions that are by commonsense standards implausible, superfluous, or incompatible with the passage. After you have chosen the best answer, blacken the corresponding space on your answer sheet.

3. More than a year ago, the city announced that police would crack down on illegally parked cars and that resources would be diverted from writing speeding tickets to ticketing illegally parked cars. But no crackdown has taken place. The police chief claims that resources have had to be diverted from writing speeding tickets to combating the city's staggering drug problem. Yet the police are still writing as many speeding tickets as ever. Therefore, the excuse about resources being tied up in fighting drug-related crime simply is not true.

The conclusion in the passage depends on the assumption that

(A) every member of the police force is qualified to work on combating the city's drug problem.
(B) drug-related crime is not as serious a problem for the city as the police chief claims it is.
(C) writing speeding tickets should be as important a priority for the city as combating drug-related crime.
(D) the police could be cracking down on illegally parked cars and combating the drug problem without having to reduce writing speeding tickets.
(E) the police cannot continue writing as many speeding tickets as ever while diverting resources to combating drug-related crime.

5. A gas tax of one cent per gallon would raise one billion dollars per year at current consumption rates. Since a tax of fifty cents per gallon would therefore raise fifty billion dollars a year, it seems a perfect way to deal with the federal budget deficit. This tax would have an additional advantage that the resulting drop in the demand for gasoline would be ecologically sound and would keep out country from being too dependent on foreign oil producers.

Which one of the following most clearly identifies an error in the author's reasoning?
(A) The author cites irrelevant data.
(B) The author relies on incorrect current consumption figures.
(C) The author makes incompatible assumptions.
(D) The author mistakes an effect for a cause.
(E) The author appeals to conscience rather than reason.

Exhibit B1. Description of the Analytical Reasoning Item Type: GRE and LSAT Versions

GRE General Test

ANALYTICAL REASONING ABILITY

Analytical reasoning questions test the ability to understand a given structure of arbitrary relationships among fictitious persons, places, things, or events, and to deduce new information from the relationships given. Each analytical reasoning group consists if (1) a set of about three to seven related statements or conditions (and sometimes other explantory material) describing a structure of relationships, and (2) three or more questions that test understanding of that structure and its implications. Although each question in a group is based on the same set of conditions, the questions are independent of one another; answering one question in a group does not depend on answering any other question.

No knowledge of formal logic or mathematics is required for solving analytical reasoning problems. Although some of the same processes of reasoning are involved in solving both analytical reasoning problems and problems in specialized fields, analytical reasoning problems can be solved using knowledge, skills, vocabulary, and computational ability (simple addition and subtraction) common to college students.

Each group of analytical reasoning questions is based on a set of conditions that establish relationships among persons, places, things, or events. These relationships are common ones such as

• temporal order (X arrived before Y but after Z),

• spatial order (City X is west of point Y and point Z),

• set membership (Professor Green serves on the committee, then Professor Brown must also serve), and

• cause and effect (Event O always causes Event R).

The conditions should be read carefully to determine the exact nature of the relationship or relationships involved. Some relationships are fixed or constant (The second house on the block belongs to P). Other relationships are variable (O must be assigned to either campsite 1 or campsite 3). Some relationships that are not given can be easily deduced from those given. (If one condition about books on a shelf specifies that book L is to the left of book Y, and another specifies that book P is to the left of book L, then it can be deduced that book P is to the left of book Y).

Law School Admission Test

ANALYTICAL REASONING

Analytical reasoning items are designed to measure the ability to understand a structure of relationships and to draw conclusions about the structure. The examinee is asked to make deductions from a set of statements, rules, or conditions that describe relationships among entities such as persons, places, things, or events. They simulate the kinds of detailed analyses of relationships that a law student must perform in solving legal problems. For example, a passage might describe four diplomats sitting around a table, following certain rules of protocol as to who can sit where. The test taker must answer questions about the implications of the given information, for example, who is sitting between diplomats X and Y.

No formal training in logic is required to answer these questions correctly. Analytical reasoning questions are intended to be answered using knowledge, skills, and reasoning ability generally expected of college students and graduates. Although each question about a passage is based on the same set of conditions, the questions are independent of one another; the answer for one question in a group does not dependent on information from any other question in the group.

The passage used for each group of questions describes relationships that are common ones such as the following:

• Assignment problems: Two parents, P and), and their children, R and S, must go to the dentist on four consecutive days, designated 1,2,3 and 4...;
• Ordering problems: X arrived before Y but after Z;

• Grouping problems: A basketball coach is trying to form a lineup from seven players--R,S,T,U,V,W, and X...and each player has a particular strength--shooting, jumping, or guarding;

• Spatial relationship problems: Country X contains six cities and each city is connected to at least one other city by a system of roads, some of which are one-way;

Careful reading and analysis are necessary to determine the exact nature of the relationships involved. Some relationships are fixed (e.g., P and R always sit at the same table). Other relationships are variable (e.g., Q must be assigned to either table 1 or table 3). Some relationships that are not stated in the conditions are implied by and can be deduced from those that are stated (e.g., if one condition about books on a shelf specifies that Book L is to the left of Book Y, and another specifies that Book P is to th e left of Book L, then it can be deduced that Book P is to the left of Book Y.)

Exhibit B2. Illustrative Analytical Reasoning Items: LSAT and GRE Versions

GRE General Test

Directions. Each question or set of questions is based on a passage or set of conditions. In answering some of the questions, it may be useful to draw a rough diagram. For each question, select the best answer choice given (one set of directions for "analytical ability" section).

In order to solve a burglary that has taken place at the Marshall's house, Detective Johnson has been instructed to search exactly five locations--the foyer, the kitchen, the living room, the guest room, the hallway, the music room, and the nursery. The foyer, kitchen, and living room are located downstairs, whereas the guest room, hallway, music room, and nursery are located upstairs. During the first visit to search for evidence, Detective Johnson has time to search exactly three locations. This search must be conducted according to the following conditions:'

The three locations searched must be neither all upstairs or all downstairs.
If the hallway is searched, then the foyer must also be searched.
If the music room is not searched, then the guest room cannot be searched.
The kitchen and the living room cannot both be searched.
The three locations searched must include the living room or the nursery, or both.

1. Which of the following is a selection of locations that conforms to the conditions for Detective Johnson's first visit?
   - (A) Foyer, guest room, hallway
   - (B) Foyer, hallway, living room
   - (C) Foyer, kitchen, living room
   - (D) Guest room, hallway, kitchen
   - (E) Guest room, music room, nursery

2. During the first visit, if the kitchen is searched, which of the following must also be searched?
   - (A) The foyer
   - (B) The guest room
   - (C) The hallway
   - (D) The music room
   - (E) The nursery

3. During the first visit, if the music room is searched, which of the following is a pair of locations that can both also be searched?
   - (A) The foyer and the hallway
   - (B) The guest room and the kitchen
   - (C) The hallway and the living room
   - (D) The kitchen and the living room
   - (E) The kitchen and the nursery

Law School Admission Test

Directions. Each group of questions is based on a set of conditions. In answering some of the questions, it may be useful to draw a rough diagram. Choose the response that most accurately and completely answers each question and blacken the corresponding space on your answer sheet.

Exactly six dogs--P,Q,R,S,T, and U--are entered in a dog show. The judge of the show awards exactly four ribbons, one for each of first, second, third, and fourth places, to four of the dogs. The information that follows is all that is available about the six dogs:

Each dog is either a greyhound or a labrador, but not both.
Two of the six dogs are female and four are male.
The judge awards ribbons to both female dogs, exactly one of which is a labrador.
Exactly one labrador wins a ribbon.
Dogs P and R place ahead of dog S, and dog S places ahead of dogs Q and T.
Dogs P and R are greyhounds.
Dogs S and U are labradors.

18. Which one of the following is a complete and accurate list of the dogs that can be greyhounds?
   - (A) P,Q
   - (B) P,R
   - (C) P,Q,R
   - (D) P,R,T
   - (E) P,Q,R,T

19. Which one of the following statements CANNOT be true?
   - (A) A female greyhound wins the second place ribbon.
   - (B) A female labrador wins the second place ribbon.
   - (C) A female labrador wins the third place ribbon.
   - (D) A male greyhound wins the fourth place ribbon.
   - (E) A female greyhound wins the fourth place ribbon.

20. Which one of the following dogs must be male?
   - (A) dog P
   - (B) dog R
   - (C) dog S
   - (D) dog T
   - (E) dog U

Exhibit C1. Descriptions of the Reading Comprehension Item Type: LSAT and GRE Versions

GRE General Test

READING COMPREHENSION

The purpose of the reading comprehension questions is to measure the ability to read with understanding, insight, and discrimination. This type of question explores the examinee's ability to analyze a written passage from several perspectives, including the ability to recognize both explicitly stated elements in the passage and assumptions underlying statements or arguments in the passage as well as the implicatins of those statements or arguments.

In each edition of the General Test, there are two relatively long reading comprehension passagen, eadch providing the basis for seven or eight questions, and two relatively short passages, each providing the basis for answering three or four questions. The four passages are drawn from four different subject areas: the humanities, the social sciences, the biological sciences, and the physical sciences.

Because the written passage upon which reading comprehension questions are based presents a sustained discussion of a particular topic, there is ample context for analyzing a variety of relationships; for example, the function of a word in relation to a larger segment of the passage, the relationships among the various ideas in the passage, or the relation of the author to his or her topic or to the audience.

There are six types of reading comprehension questions. These types focus on

- the main idea or primary purpose of the passage;
- information explicitly stated in the passage;
- information or ideas implied or suggested by the author;
- possible application of the author's ideas to other situations;
- the author's logic, reasoning, or persuasive techniques; and
- the tone of the passage or the author's attitude as it is revealed in the passage.

Law School Admission Test

READING COMPREHENSION

The purpose of reading comprehension questions is to measure your ability to read with understanding and insight, examples of lengthy and complex materials similar to those commonly encountered in law school work. The Reading Comprehension section of the test consists of four passages, each approximately 450 words long, followed by five to eight questions that test reading and reasoning abilities. Passages for reading comprehension items offer variety in styles of writing and draw from a wide variety of subjects--including the humanities, the social sciences, the physical sciences, ethics, philoslophy, and the law.

Reading comprehension questions require test takers to read carefully and accurately, to determine the relationships among the various parts of the passage, and to draw reasonable inferences from the material in the passage. The questions may ask about:

- the main idea or primary purpose of the passage;
- the meaning or purpose of words or phrases used in the passage;
- information or ideas that can be inferred from the passage;
- how the author makes his or her points; and
- the tone of the passage or the author's attitude as it is revealed in the language used.

Exhibit C2a. Illustrative GRE Reading Comprehension Items (Abridged Options)

Directions: Each passage in this group is followed by questions based on its content. After reading a passage, choose the best answer to each question. Answer all questions following a passage on the basis of what is stated or implied in the passage.

The more that is discovered about the intricate organization of the nervous system, the more it seems remarkable that genes can successfully specify the development of that system. Human genes contain too little information even to specify which hemisphere of the brain each of a human's 10^{11} neurons should occupy, let alone the hundreds of connections that each neuron makes. For such reasons , we can assume that there must be an important random factor in neural development, and in particular, that errors must and do occur in the development of all normal brains.

The most vivid expression of such errors occurs in genetically identical organisms. Even when reared under the same conditions, isogenic organisms are rarely exact copies of one another, and their differences have revealed much about the random variations that result from an organism's limitied supply of genetic information. In isogenic Daphniae, for example, even though the position, size, and branching pattern of each optic neuron are remarkably constant, there is some variability in connectivity, and the number of synapses varies greatly. This variability is probably the reult of random scatter beyond the resolution of genetic control and is best termed "imprecision," since its converse, the degree of clustering about a mean, is conventionally called "precision."

Imprecision should be distinguished from developmental mistakes: wrongly migrated neurons, incorrect connections, and the like. To use a computer analogy, minor rounding-off errors occur universally and are analogous to imprecision, but occasionally a binary digit is incorrectly transmitted, perhaps ruining a calculation, and this incorrect transmission is analogous to a developmental mistake. Thus imprecision is a form of inaccuracy inherent within the limits of design, but mistakes are forms of gross fallibility.

Both imprecision and gross fallibility can plausibly be blamed on the insufficiency of genetic information, since either cound be reduced by adding more information. It is universally accepted amonjg information theorists that codes and languages can be made mistake-resistant by incorporating redundancy. However, since the amount of space available in any information system is limited, increased redundancy results in decreased precision. For example, , when written in English, "three point oen four two," can be understood correctly even though a typographical error has occurred. MOre precision could be gained, however, if those 24 spaces were filled with Arabic numerals; then could be expressed to 23 significant digits;, although any error would significantly change the meaning. There exists a trade-off; the more precisely a system is specified, using a given limited amount of information, the greater the danger of gross mistakes;. The overall scheme by which genetic information is rationed out in an organism, therefore, must involve a compromise between two conflicting priorities: precision and the avoidance of gross mistakes.

1. Which of the following best expresses the main idea of the passage?

(A) Although studies of isogenic organisms have shown that all organisms are subject to developmental variations, there is still scientific debate over the exact causes of these variations.

(B) Because of limitations on the amount of information contained in the genes of organisms, developing nervous systems are subject to two basic kinds of error, the likelihood of one of which is reduced only when the likelihood of the other is increased.

(C) The complexity of an organism's genetic information meeans that much of the unusual variation that occurs among organism can best be explained as the result of developmental mistakes.

(D) New findings about the nature of the genetic control of neural development support the work of some scientists who argue that the computer is an extremely useful model for the understanding of the nervous system.

(E) The major discovery made by scientists studying the genetic control of neural development is that both imprecision and gross developmental error can be traced to specific types of mutations in genes.

2. According to the passage, one of the reasons it has been assumed that there is an important random element in human neural development is that

(A) genes cannot specify certain types of developmental processes as well as they can others (FOUR ADDITIONAL OPTIONS)

3. The author suggests which of the following about the findings of information theorists? FIVE OPTIONS

4. According to the passage, of the following aspects of the optic neurons of isogenic Daphniae, which varies the most? FIVE OPTIONS

5. Which of the following best describes the organization of the first paragraph? FIVE OPTIONS

6. The author uses all of the following to clarify the distinction between imprecision and gross mistakes in neural development EXCEPT
FIVE OPTIONS

7. Which of the following can be inferred from the passage about the genetic information of Diaphniae? FIVE OPTIONS

Exhibit C2b. Illustrative LSAT Reading Comprehension Passage and Questions (Abridged Options)

__Directions.__ Each passage in this section is followed by a group of questions to be answered on the basis of what is __stated__ or __implied__ in the passage. For some of the questions, more than one of the choices could conceivably answer the question. However, you are to choose the __best__ answer; that is, the response that most accurately and completely answers the question, and blacken the corresponding space on your answer sheet.

Three basic adaptive responses--regulatory, acclimatory, and developmental--may occur in organisms as they react to changing environmental conditions. In all three, adjustment of biological features (morphological adjustment) or of their use (functional adjustment) may occur. Regulatory responses involve rapid changes in the organism's use of its physiological apparatus--increasing or decreasing the rates of various processes, for example. Acclimation involves morphological change--thickening of fur or red blood cell proliferation--which alters physiology itself. Such structural changes requir more time than regulatory response changes. Regulatory and acclimatory responses are both reversible.

Developmental responses, however, are usually permanent and irreversible; they become fixed in the course of the individual's development in response to environmental conditions at the time the response occurs. One such response occurs in many kinds of water bugs. Most water-bug species inhabiting small lakes and ponds have two generations per year. the first hatches during the spring, reproduces during the summer, then dies. The eggs laid in the summer hatch and develop into adults in late summer. They live over the winter before breeding in early spring. Individuals in the second (overwintering) generation have fully developed wings and leave the water in atuumn to overwinter in forests, returning in spring to small bodies of water to lay eggs. Their wings are absolutely necessary for this seasonal dispersal. The summer (early) generation, in contrast, is usually dimorphic--some individuals have normal functional (macropterous) wings; others have much-reduced (micropterous) wings of no use for flight. The summer generation's dimorphism is a compromise strategy, for these individuals usually do not leave the ponds and thus generally have no use for fully developed wings. But small ponds occasionally dry up during the summer, forcing the water bugs to search for new habitats, an eventuality that macropterous individuals are well adapted to meet.

The dimorphism of micropterous and macropterous individuals in the summer generation expresses developmental flexibility; it is not genetically determined. The individual's wing form is environmentally determined by the temperature to which developing eggs are exposed prior to their being laid. Eggs maintained in a warm environment always produce bugs with normal wings, but exposure to cold produces micropterous individuals. Eggs producing the overwintering brood are all formed during late summer's warm temperatures. Hence, all individuals in the overwintering brood have normal wings. Eggs laid by the overwintering adults in the spring, which develop into the summer generation of adults, are formed in early autumn and early spring. Those eggs formed in autumn are exposed to cold winter temperatures, and thus produce micropterous adults in the summer generation. Those formed during the spring are never exposed to cold temperatures and thus yield individuals with normal wings. Adult water bugs of the overwintering generation, brought into the laboratory during the cold months and kept warm, produce only macropterous offspring.

1. The primary purpose of the passage is to

(A) illustrate an organism's functional adaptive response to changing environmental conditions

(B) prove that organisms can exhibit three basic adaptive responses to changing environmental conditions

(C) explain the differences in form and function between micropterous and macropterous water bugs and analyze the effect of environmental changes on each

(D) discuss three different types of adaptive responses and provide an example that explains how one of those types of responses works

(E) contrast acclimatory responses with developmental responses and suggest an explanation for the evolutionary purposes of these two responses in changing environmental conditions

2. The passage supplies information to suggest that which one of the following would happen if a pond inhabited by water bugs were to dry up in June?

(A) The number of developmental responses among the water-bug population would decreas (FOUR ADDITIONAL OPTIONS)

3. It can be inferred from the passage that if the winter months of a particular year were unusually warm, the (FIVE OPTIONS)

4. According to the passage, the dimporphic wing structure of the summer generation occurs because (FIVE OPTIONS)

5. It can be inferred from the passage that which one of the following is an example of a regulatory response? (FIVE OPTIONS)

6. According to the passage, the generation of water bugs hatching during the summer is likely to (FIVE OPTIONS)

7. The author mentions laboratory experiments with adult water bugs (lines 63 to 66) in order to illustrate which one of the following? (FIVE OPTIONS)

8. Which one of the following best describes the organization of the passage? (FIVE OPTIONS)

Based on the strong surface similarities evident in these exhibits, it seems clear that although items for the two tests are developed to somewhat different specifications (subject matter, type of situations depicted, relative emphasis on particular subskills, and so on) logical reasoning, analytical reasoning, and reading comprehension items from the LSAT and the GRE can be thought of as measuring generally similar skills and abilities, and are in fact quite similar with respect to general characteristics, subskills targeted for assessment, and so on.

Differences between the two tests in the way in which these particular item types are presented to examinees, the number of items included, amount of time devoted to each item type, and so on, are indicated in Table 2.2.

Table 2.2.

General Organization of Tasks in the LSAT and in the GRE General Test, Respectively

LSAT sections	No. items	Time
Logical Reasoning	24	35 minutes
Reading Comprehension	28	35 minutes
Analytical Reasoning	24	35 minutes
Logical Reasoning	25	35 minutes
Variable section	Variable	35 minutes
TOTAL SCORE	101	
GRE sections		
Verbal section 1	38	30 minutes
Item type		
Analogy	9	
Sentence Completion	7	
Reading Comprehension	11	
Antonym	11	
Verbal section 2	38	30 minutes
Analogy	9	
Sentence Completion	7	
Reading Comprehension	11	
Antonym	11	
Quantitative section 1		30 minutes
Item type	30	
Quantitative comparison	15	
Regular mathematics	10	
Data interpretation	5	
Quantitative section 2	30	30 minutes
Quantitative comparison	15	
Regular mathematics	10	
Data interpretation	5	
Analytical section 1	25	30 minutes
Item type		
Logical reasoning	6	
Analytical reasoning	19	
Analytical section 2	25	
Item type	Number of items	
Logical reasoning	6	
Analytical reasoning	19	
Variable section (not scored)	Variable	30 minutes
Total number of test items	186	

In forms of the LSAT, including the forms involved in the present study, logical reasoning, reading comprehension, and analytical reasoning items are presented in four separately timed 35-minute sections, each of which is homogeneous with respect to item type. Two of these sections are made up of logical reasoning items. A fifth (variable) section is included for purposes of test development. Performance on the items in the variable section does not affect a candidate's LSAT score.

The ordering of LSAT sections differs from form to form, and across operational versions of given generic forms.

The GRE includes a total of seven separately timed sections, each with a 30-minute time limit, one of which is variable. In a given version of the test, sections are presented in such a way as to require examinees to work sequentially through two different sets of tasks from the respective ability domains—for example, a 30-minute verbal set, a 30-minute quantitative set, a 30-minute analytical set, a second quantitative set, a second analytical set, a second verbal set, followed by a variable set.

The two sections of GRE items that constitute a given ability measure, although developed to comparable specifications, are not designed to be parallel with respect to validity, difficulty level, and so on.

SECTION 3. ASSESSING THE DIMENSIONALITY OF THE LSAT

This section describes factor analyses that were undertaken

(a) to explore the extent to which the Law School Admission Test—made up of logical reasoning, analytical reasoning, and reading comprehension item types, organized as indicated in Table 3.1—measures more than one identifiable aspect of reasoning, and

(b) to help clarify the aspects of reasoning tapped by the three LSAT item types.

Table 3.1.

Organization of the LSAT

LSAT Sections	Number of Items	Acronym	Time
Logical reasoning	24	LR24	35 minutes
Reading comprehension	28	RC28	35 minutes
Analytical reasoning	24	AR24	35 minutes
Logical reasoning	25	LR25	35 minutes
Variable section	Variable		35 minutes
TOTAL SCORE	101	LSAT	

Extensive research findings in the GRE context indicate that the GRE versions of the item types used in the LSAT are tapping psychometrically distinguishable aspects of reasoning. And, as we have seen, published descriptions of the three LSAT item types are very similar to the published descriptions of their GRE counterparts, and actual items of the three types from the LSAT appear to be quite similar, with respect to surface characteristics, to those appearing in the GRE General Test.

General Analytical Rationale

With respect to the dimensionality issue, GRE research findings indicate relatively clearly that

- the aspects of reasoning measured by the logical reasoning item type are closely related to those tapped by the reading comprehension item type, but questions as to the ability-domain classification of the two item types have not been resolved, but

- the aspects of reasoning tapped by the analytical reasoning item type appear to be psychometrically distinguishable from those tapped by reading comprehension items and logical reasoning items—albeit not as clearly so for logical reasoning as for reading comprehension item types.

Judging from correlational findings involving LSAT item-type section scores, available for two LSAT forms involved in the present study (from unpublished ETS internal analyses), patterns of relationships involving the respective item types in the LSAT context, suggest interpretive inferences that are similar to those outlined above for the GRE counterparts. The LSAT findings are summarized in Table 3.2.

Table 3.2.
Observed Correlations and Correlations After Correction for Attenuation: LSAT Section Scores, June 1991 and October 1991 Forms of the LSAT*

	LSAT Form							
	June 1991				October 1991			
Section	LR25	LR24	RC28	AR24	LR25	LR24	RC28	AR24
LR25	(.78)	.97	.91	.71	(.77)	.96	.89	.68
LR24	.76	(.79)	.89	.72	.74	(.77)	.87	.64
RC28	.72	.71	(.80)	.63	.69	.68	(.79)	.59
AR24	.55	.56	.50	(.77)	.52	.49	.46	(.76)

Note. Observed correlations are shown below the diagonal: corrected correlations are shown above the diagonal; diagonal elements are estimated KR-20 reliabilities, used to correct the correlations.

* Data from unpublished ETS internal test analyses for forms of the LSAT used in the present study; KR-20, internal-consistency estimates.

The pattern of correlations shown in Table 3.2 is particularly instructive because corrected correlations for the two LSAT logical reasoning sections (r = .97 and r = .96)—coefficients approaching 1.00, consistent with the fact that both sections are intended to measure the same underlying skills and abilities—can be contrasted directly with corrected correlations between the logical reasoning section scores and scores for sections involving two different item types, namely,

(a) score for the reading comprehension item type, tapping abilities that are quite similar to those tapped by the LSAT logical reasoning sections (r's ranging between .87 and .91, but lower than those for the two logical reasoning sections), and

(b) score for the analytical reasoning item type, tapping abilities that do not tend to overlap markedly with those tapped by logical reasoning (indicated by corrected correlations ranging between .64 and .72) or by reading comprehension, for which corrected correlations with analytical reasoning were .59 and .63.

It can also be seen that correlations between LSAT reading comprehension and analytical reasoning (r = .59 and r = .63) are lower than those between LSAT logical reasoning and analytical reasoning (r's between .64 and .72).

The marked overlap between reading comprehension and logical reasoning reflected in the corrected correlations in Table 3.2, suggests that two factors might explain the basic underlying patterns of intercorrelations, namely, one factor defined by logical reasoning and reading comprehension parcels, and a second factor defined exclusively by analytical reasoning parcels.

At the same time, the correlational gradient evident in Table 3.2, indicates limited differentiation of logical reasoning from reading comprehension, as well as relatively clear differentiation of both of these item types from the analytical reasoning item type. Accordingly, it was considered useful to evaluate a three-factor solution in order to leave logical reasoning and reading comprehension item types free to demonstrate their (limited) distinctiveness by defining separate factors, along with a factor defined exclusively by the LSAT analytical reasoning item type.[6]

Thus, based on indirect evidence from the GRE context and direct evidence from the LSAT context, it was assumed a priori that at least two common factors would be needed to explain intercorrelations of scores on LSAT item-type parcels (sets of four to six items of the respective types), and that one of the factors would be defined primarily by the analytical reasoning item type.

Study Sample and Data

A file containing records for more than 74,000 LSAT examinees tested in either June 1991 or October 1991 LSAT administrations was provided by Law School Admission Services (LSAS). Each record contained test data, including rights-scored LSAT items, responses to several background questions (gender, ethnic group membership, date of birth, and so on). More comprehensive information (for example, undergraduate school, level of degree attained, undergraduate grades, undergraduate major field) was available only for examinees who participated in the Law School Data Assembly Service (for example, LSAT, 1991).

Sampling from the LSAS file was conducted in two phases, one of which was concerned simply with obtaining a general sample of LSAT examinees for inclusion in the study; the other sampling phase was concerned with identifying LSAT examinees with GRE records.

In the first phase, LSAT records from the general LSAS file were selected for a spaced sample of 7,385 examinees—all of whom were native-English speakers, with unflagged records, who did not report a handicapping condition.

The second phase required matching records in the LSAS file with records in GRE files in order to identify a joint sample. Direct evidence as to either the extent of "overlapping" of LSAT and GRE test-taking populations, or the patterns of test-taking involved—with respect to sequence, interval between tests, and so on—was not available.

File matching involving GRE administrations subsequent to the October 1991 LSAT administration was necessarily limited by study timing. It was decided to use an iterative approach, in which the LSAS file would be matched successively with

---

[6] In GRE factor studies, without exception, item-level, parcel-level, or other scores involving these three item types have been analyzed in the context of scores on a more comprehensive "item-type battery," rather than in analyses focussed exclusively on these three item types, as in the present instance.

GRE files, beginning with the file for the December 1991 GRE administration and working backward until a "sufficient sample" was identified.

Using this approach, the LSAT file was matched across files for all GRE administrations between October 1988 and December 1991, resulting in the identification of a total of 5,281 LSAT examinees (from the two LSAT administrations) who had GRE records as well. Records in the joint file for examinees identified in the LSAT file as nonnative-English speakers, or who reported some form of handicapping condition (for example, deafness, impaired vision), and records that were "flagged," indicating exceptional test-related circumstances, were excluded from the joint sample, as was done for the general LSAT sample.[7] After purging this selected file, the sample remaining (N = 4,447) was judged to be adequate for study purposes.

The analyses reported in this section are based exclusively on LSAT data for the general sample of LSAT examinees (N = 7,385) and LSAT data for the selected sample (N = 4,447) of LSAT examinees who also took the GRE. The distribution of these examinees according to LSAT form taken, and descriptive statistics for selected LSAT-related variables, by test form, are shown in Table 3.3. For added perspective, means and standard deviations of scaled GRE scores, available only for the selected sample, are shown as well.

---

[7] Generally speaking, unless otherwise indicated, data in the LSAT file were used for purposes of describing and classifying examinees. Questions regarding the "reliability" of self-report regarding stable personal attributes, of considerable interest generally, were not at issue in the present study.

Table 3.3.

Descriptive Statistics for General and Selected LSAT Samples, by LSAT Form Taken: Designated Variables

	Sample							
	General				Selected			
	June		October		June		October	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
LSAT	151.7	9.8	152.1	8.8	154.3	10.2	154.7	9.0
LR24	14.8	4.5	14.9	4.3	15.6	4.6	16.0	4.3
LR25	14.3	4.5	16.3	4.3	15.5	4.7	17.3	4.4
RC28	17.8	4.9	17.1	4.9	19.4	4.9	18.9	5.0
AR24	11.4	4.6	13.3	4.4	12.1	4.7	13.9	4.4
LSDAS* (%)	58.5		62.7		55.1		65.4	
Female (%)	43.1		45.2		51.2		50.3	
GRE Verbal		n.a.			551	112	562	108
GRE Quantitative		n.a.			567	130	578	124
GRE Analytical		n.a.			602	125	618	117

Note. Based on unpublished internal analyses at ETS, the LSAT scaled score mean for the June 1991 administration was 151.29 and the standard deviation was 9.98; for the October administration, corresponding values were 151.74 and 8.81.

* Estimated percentage participating in the Law School Data Assembly Service (e.g., LSAT, 1991).

It can be seen in the table that LSAT scaled-score means for the selected (LSAT/GRE) sample were higher than those for the general samples (by almost three-tenths of a standard deviation). It is noteworthy, however, that standard deviations for LSAT-related variables in the selected sample are comparable to those in the general sample.

As a matter of incidental interest in the present context, on the GRE side, estimated percentile-rank equivalents of the three GRE means (ETS, 1992, Table IA), place the selected LSAT/GRE sample at approximately the 68th percentile for GRE verbal, the 50th percentile for GRE quantitative, and the 66th percentile for GRE analytical.

Analytical Approach

To assess factors underlying LSAT performance, analyses were made of intercorrelations among parcels of LSAT items, as defined below. The 101 items in the LSAT were divided into 20 parcels-sets of logical reasoning, reading comprehension, and analytical reasoning items, each made up of four to six items of a particular type—distributed by test section as shown in Table 3.4.

Table 3.4.
General Definitional Characteristics of Parcels of LSAT Items

Section Acronym	Parcel Designation (Number items per parcel)	Section Description
LR25	L51 L52 L53 L54 L55 (5) (5) (5) (5) (5)	(25-item LR section)
LR24	L61 L62 L63 L64 L65 (5) (5) (5) (5) (4)	(24-item LR section)
RC28	R1 R2 R3 R4 R5 (6) (6) (6) (5) (5)	(28-item RC section)
AR24	AR1 AR2 AR3 AR4 AR5 (5) (5) (5) (5) (4)	(24-item AR section)

Items were assigned to parcels in such a way that each parcel contained a mix of items that was roughly balanced with respect to item sequence within a section.

- For example, the first parcel from the 25-item LR section was composed of items 1, 6, 11, 16, and 21, as was the first parcel from the 24-item LR section and the 24-item AR section; for the 28-item RC section, the first parcel was composed of items 1, 6, 11, 16, 21, and 28.

This pattern of assigning items to parcels was designed to attain rough balance across parcels with respect to difficulty (assuming that items appearing later in a test section tend to be more difficult than those appearing earlier in a section), and to reduce the likelihood of interpretive complications associated with possible end-of-section effects (for example, speededness, fatigue).

Analytic Procedure and Related Findings

The factor analytic approach employed was exploratory in nature, but with respect to both design and interpretation the approach used was guided by evidence, assumptions, and lines of reasoning developed in detail above—that is, from the outset, attention was focussed on the exploration of both two-factor and three-factor solutions for the data under consideration. Both orthogonal (varimax) and oblique (direct oblimin) rotations, involving the corresponding principal components, were used to obtain two-factor and three-factor solutions. In all instances, correlation matrices with unity in the diagonal were employed.

Two- and three-factor solutions were obtained (a) for the two June-form subsamples (that is, the general sample and the selected sample), and the corresponding October-form subsamples. Analyses in these four subsamples involved raw scores on item-type parcels from either the June form or the October form. Two additional analyses were made, one involving pooled general-sample data for the June and October forms, and the other involving pooled selected-sample data for both test forms. To summarize succinctly, the several analysis samples involved were as indicated below.

	Form Classification		
Sample Classification	June Form (N)	October Form (N)	Combined Forms (N)
LSAT, general	2,429	4,956	7,385
LSAT, selected	1,389	3,058	4,447

Analyses involving either the June form or the October form were based on raw parcel scores. However, for the two combined-forms analyses (involving data for the two different LSAT forms), scores on parcels (and other raw scores) were z-scaled within the respective form-samples, prior to computing the intercorrelations. The resulting matrices reflect the pooled, within-form relationships that are of primary interest in assessing factors that reflect the internal structure of the LSAT.

Focus on pooled, within-form correlations is needed because of the possibility of interpretive complications if raw scores (e.g., number-right parcel scores or section scores) from two different forms of the LSAT (GRE, or other similarly constructed standardized admission test) are combined for analysis. The possibility of interpretive complications arises from the fact that sets of items of particular types are not designed to be, and typically are not "parallel" across test forms—that is, they tend to differ in difficulty, validity, and so on.

The exploratory pooling approach outlined above, involves an assumption that test items of the same type and position, in successive test forms developed under controlled conditions to the same general specifications, are statistically interchangeable for study purposes; also that internal construct validity tends to be invariant across test forms. Relatively few studies appear to have been conducted explicitly to assess consistency of internal construct validity across successive test forms, and none appears to have analyzed pooled within-form correlation matrices. For example, Dorans and Lawrence (1987) found reasonable consistency in structure across six SAT forms, but were not directly concerned with questions regarding pooling data across these forms.

In essence there appears to be no direct precedent for the exploratory approach employed in this study. Thus, the across-form analyses involving pooled within-form LSAT correlations (based on z-scaled parcels from only two administrations) provide perspective for evaluating outcomes of similarly designed analyses of pooled within-form GRE correlations, involving 21 different forms of the GRE, to be considered in Section 4.

Replication in the six analysis-samples outlined above was designed to permit assessment of the consistency of factor outcomes for two different forms of the LSAT: consistency in outcomes involving the selected LSAT form-samples and outcomes involving the general LSAT form-samples, if present, would suggest that data for the selected LSAT sample provide a sound basis for inferences about LSAT factor structure—and for evaluation of correlations between the LSAT and the GRE, to be reported later. Similarly, consistency in outcomes for correlations involving z-scaled parcels pooled across form-samples and outcomes of separate analyses for the June and October form samples would suggest the viability of the pooling approach involved.

Generally speaking, outcomes of both two-factor and three-factor solutions were in all major respects, consistent across the six analysis-samples.

With respect to the principal components phase of the analysis, summarized selectively in Table 3.5, for example, it is evident that in each analysis-sample,

- the first two principal components, with eigenvalues greater than one, accounted for some 45 percent to 50 percent of total variance;
- adding the third component, with eigenvalues in the .9 range, resulted in solutions involving between 50 percent and 54 percent of the total variance.

Table 3.5.

Summary of Principal Components Extractions in
Six LSAT Analysis-Samples: Three Factors

	Eigenvalues			Percent total variance				
Analysis-Sample	PC1	PC2	PC3	PC1	PC2	(1+2)	PC3	(Cum)
General, June	7.94	1.64	.86	39.7	8.2	(47.9)	4.3	(51.5)
General, October	7.45	1.63	.96	37.3	8.1	(45.4)	4.8	(50.2)
General, Combined	7.61	1.62	.91	38.0	8.1	(45.4)	4.6	(50.7)
Selected, June	8.49	1.56	.83	42.4	7.8	(50.2)	4.2	(54.4)
Selected, October	7.90	1.65	.92	39.5	8.2	(47.8)	4.6	(52.4)
Selected, Combined	8.08	1.61	.88	40.4	8.1	(48.4)	4.4	(52.8)

With respect to the rotational phase of the analysis, involving orthogonal (varimax) and oblique (direct oblimin) rotations to two-factor and three-factor solutions in each analysis-sample, findings bearing on factor identification and structure were also consistent from sample to sample. Accordingly, detailed findings are presented and discussed only for the general, combined-forms sample (N - 7,385).

Attention is directed first to the orthogonal factors identified in the two-factor and three-factor varimax rotations, shown in Table 3.6. The resulting factors, in every instance, based on the observed loadings of parcels, are strongly and cleanly identified.

Table 3.6.

Varimax Factors for Two-Factor and Three-Factor Solutions in the General LSAT, Combined-Forms Analysis-Sample (N = 7,385)

Rotated factor matrix

Parcel	Factor 1	Factor 2
L51	.59	.23
L52	.62	.24
L53	.59	.19
L54	.62	.25
L55	.61	.23
L61	.62	.21
L62	.60	.23
L63	.65	.25
L64	.63	.24
L65	.54	.25
R1	.68	.16
R2	.64	.13
R3	.67	.15
R4	.62	.13
R5	.63	.12
AR1	.24	.71
AR2	.21	.69
AR3	.20	.71
AR4	.25	.74
AR5	.19	.72

Note. Varimax converged in 3 iterations.

Rotated factor matrix

Parcel	Factor 1	Factor 2	Factor 3
L51	.62	.18	.21
L52	.59	.20	.27
L53	.64	.13	.17
L54	.60	.21	.27
L55	.59	.19	.26
L61	.61	.17	.25
L62	.59	.19	.25
L63	.60	.21	.31
L64	.61	.20	.27
L65	.59	.20	.16
R1	.31	.17	.68
R2	.29	.14	.66
R3	.33	.16	64
R4	.28	.14	.64
R5	.25	.14	.68
AR1	.20	.70	.17
AR2	.21	.68	.11
AR3	.17	.71	.15
AR4	.23	.73	.16
AR5	.20	.71	.10

Note. Varimax converged in 6 iterations.

In the two-factor solution shown in the table (as in the corresponding solutions for the remaining analysis-samples) the first factor to emerge is strongly identified by the logical reasoning and reading comprehension parcels, and the second is equally strongly identified by the analytical reasoning parcels. This is consistent with expectation based on the assumptions and lines of reasoning outlined above.

In the three-factor solution, factors corresponding to the three item types emerge quite clearly: the first factor to emerge is identified strongly by the logical reasoning parcels, the second is strongly identified by the five analytical reasoning parcels, and the third is equally strongly identified by the five reading comprehension parcels.

Factor pattern coefficients for parcels in two-factor and three-factor direct oblimin rotations in the combined-forms sample are shown in Table 3.7. Again, oblimin outcomes for all other analysis-samples were quite consistent with those selected here for discussion.

Table 3.7.

Pattern Coefficients for LSAT Parcels on Oblimin Factors: Two-Factor and Three-Factor Solutions in the General LSAT, Combined-Forms Sample (N = 7,385)

	Pattern matrix	
Parcel	Factor 1	Factor 2
L51	.61	.05
L52	.63	.06
L53	.62	.01
L54	.63	.07
L55	.62	.05
L61	.64	.03
L62	.62	.06
L63	.66	.06
L64	.64	.06
L65	.54	.10
R1	.72	-.05
R2	.69	-.08
R3	.72	-.06
R4	.67	-.07
R5	.68	-.08
AR1	.04	.72
AR2	.01	.72
AR3	-.01	.74
AR4	.05	.75
AR5	-.02	.76

Note. Oblimin converged in 4 iterations; "delta" set at 0.

	Pattern matrix		
Parcel	Factor 1	Factor 2	Factor 3
L51	.69	-.00	-.03
L52	.62	.02	.06
L53	.76	-.06	-.08
L54	.64	.03	.06
L55	.63	.01	.05
L61	.66	-.02	.04
L62	.64	.01	.03
L63	.61	.03	.10
L64	.65	.02	.05
L65	.67	.04	-.08
R1	.03	.03	.74
R2	.02	-.01	.72
R3	.08	.00	.68
R4	.01	.00	.71
R5	-.05	.00	.76
AR1	-.01	.73	.04
AR2	.04	.72	-.04
AR3	-.05	.76	.03
AR4	.02	.76	.02
AR5	.01	.76	-.05

Note. Oblimin converged in 5 iterations; "delta" set at 0.

Outcomes with respect to factor correlations, in both the two-factor and three-factor oblimin solutions, were quite consistent across analysis samples, as indicated in Table 3.8. For perspective, it is useful to compare the factor correlations shown in the Table 3.8, with observed correlations for subscores based on the item types that principally define the respective factors. To this end, subscores were computed for the 49 logical reasoning items (LR49) and for the 77 logical reasoning and reading comprehension items combined (VBL77), and intercorrelations were computed involving these two subscores, the 28-item reading comprehension subscore (RC28), and the 24-item analytical reasoning subscore (AR24).

To summarize briefly:

- The VBL77 subscore (LR+RC) correlated .57 with the analytical reasoning subscore—versus .54 for F1,F2 in the table.
- The consolidated logical reasoning subscore (LR49) correlated .73 with that for reading comprehension (RC28), and .57 with the analytical reasoning subscore (AR24)—versus .67 and .54 for F1,F3 and F1,F2 in the table.
- The analytical reasoning subscore correlated .47 with the reading comprehension subscore—versus .44 for F2,F3 in the table.

Table 3.8.

Summary of Factor Correlations for Two-Factor and Three-Factor Oblimin Solutions in the Six Analysis Samples

	Factor Correlations			
	Two Factors*	Three Factors#		
Analysis Sample	F1,F2	F1,F2	F1,F3	F2,F3
General, June	.54	.55	.70	.41
General, October	.54	.56	.66	.41
General, Combined	.54	.53	.67	.43
Selected, June	.56	.56	.70	.46
Selected, October	.58	.57	.67	.44
Selected, Combined	.54	.54	.67	.44

* Based on primary identification by LSAT parcels, F1 reflects primarily logical reasoning and reading comprehension, and F2 reflects analytical reasoning.

# F1 reflects primarily logical reasoning, F2 reflects analytical reasoning, and F3 reflects reading comprehension.

The findings that have been reported in this section, suggest that the three item types that are included in the LSAT are measuring psychometrically distinguishable aspects of reasoning ability, namely, aspects of general reasoning ability, tapped in slightly differing ways[8] by reading comprehension and logical reasoning items, versus aspects of more formal, deductive reasoning ability, tapped by the analytical reasoning item type.

Analyses involving these three item types in the GRE General Test are reported in detail in the following section. Among other things, it will be seen that outcomes of the separate analysis of this subset of GRE items, generally parallel those reported above for the corresponding LSAT items.

SECTION 4. ANALYSIS OF FACTORS UNDERLYING PERFORMANCE ON THE GRE GENERAL TEST IN THE SELECTED LSAT/GRE SAMPLE

The analyses reported in Section 3 were concerned with the internal structure of the LSAT. The analyses reported in this section have to do with factors underlying performance on (a) the full array of GRE item types, and (b) a subset of the full array (logical reasoning, analytical reasoning, and reading comprehension items), the latter paralleling analyses of the corresponding LSAT item types, reported in Section 3.

Before describing these analyses, a brief review is provided of (a) the make-up of the GRE General Test, with respect to item-type composition and the organization of item-type sections, time allocated, and so on, (b) the unique GRE data-set involved in these analyses, and (c) procedures employed in aggregating data across multiple forms of the GRE are described.

The GRE Data Set

As shown earlier (see Table 2.1, Section 2), the GRE General Test involves six operational sections—two designed to measure "verbal ability," two designed to measure "quantitative ability," and two designed to measure "analytical ability." Recall that each GRE verbal section includes a total of 38 reading comprehension, analogy, antonym, and sentence-completion items; each GRE quantitative section includes a total of 25 quantitative comparison, regular mathematics, and data interpretation items[9]; and each GRE analytical section includes a total of 25 analytical reasoning and logical reasoning item types.

---

[8] The distinction between what is measured by logical reasoning items and reading comprehension items may to some extent be "blurred" by the heavy reading comprehension requirements in the logical reasoning item type here under consideration. To the extent that it is possible to do so, developing logical reasoning item types with more limited reading comprehension demands should help to clarify distinctions between these closely related aspects of general reasoning ability.

To review briefly, as reported in detail in Section 3, a file containing records for more than 74,000 LSAT examinees tested in either June 1991 or October 1991, was matched with GRE files spanning test administrations between October 1988 and December 1991. Some 5,281 LSAT examinees were identified in this matching process as having also taken the GRE.

After purging the file of records for nonnative-English speakers and examinees with "flagged" LSAT records, a total of 4,447 records was available in the joint LSAT/GRE file. As shown in the preceding section, these individuals were relatively highly selected with respect to performance on the GRE General Test as well as with respect to performance on the LSAT.

Aggregating Data Across GRE Forms

Collectively, sample members took 33 different versions (editions) of the GRE General Test, 21 of which were generic forms of the test—that is, forms differing completely (with respect to content) from previous GRE forms. The other 12 versions were content-equivalent versions of one or another of the 21 generic forms—that is, versions with essentially the same item content, though not necessarily in the same format (for example, editions for examinees with handicaps). These modified versions were treated as comparable to the corresponding generic forms for study purposes. Table 4.1 shows the number of examinees taking each of the 21 generic forms of the GRE, by form of the LSAT taken. Table 4.2 shows the distribution of the 37 general or special GRE test administrations involved.

---

[9] The quantitative comparison item type and the data interpretation item type were used, though not concurrently, in several versions of the LSAT prior to 1982.

Table 4.1. Distribution of GRE Forms, by LSAT Form Taken

GRE form*	LSAT form June	LSAT form Oct	
41 Row Column	23 25.6 1.7	67 74.4 2.2	90 2.0 (Col %)
42	69 28.9 5.0	170 71.1 5.6	239 5.4
43	16 45.7 1.2	19 54.3 .6	35 .8
44	33 25.6 2.4	96 74.4 3.1	129 2.9
45	27 35.5 1.9	49 64.5 1.6	76 1.7
46	122 25.9 8.8	349 74.1 11.4	471 10.6
47	39 27.5 2.8	103 72.5 3.4	142 3.2
48	40 41.7 2.9	56 58.3 1.8	96 2.2
49	42 29.6 3.0	100 70.4 3.3	142 3.2
50	63 28.1 4.5	161 71.9 5.3	224 5.0

* GRE form numbers are as assigned for study identification.

Table 4.1, concluded

GRE form*	LSAT form June	LSAT form Oct	
51	38 25.7 2.7	110 74.3 3.6	148 3.3
52	63 13.5 4.5	139 8.7 4.5	202 4.5
53	187 41.4 13.5	265 58.6 8.7	452 10.2
54	61 31.6 4.4	132 68.4 4.3	193 4.3
55	110 29.5 7.9	263 70.5 8.6	373 8.4
56	79 36.2 5.7	139 63.8 4.5	218 4.9
57	76 31.9 5.5	162 68.1 5.3	238 5.4
58	178 26.5 12.8	493 73.5 16.1	671 15.1
59	30 28.6 2.2	75 71.4 2.5	105 2.4
60	49 43.0 3.5	65 57.0 2.1	114 2.6
61	44 49.4 3.2	45 50.6 1.5	89 2.0
Column Total	1389 31.2	3058 68.8	4447 100.0

Table 4.2. Distribution of the LSAT/GRE Sample, by GRE Testing Date

	GRE administration Year				
Month	1988	1989	1990	1991	Total
Jan		4	4	7	15 .3
Feb		93	119	184	396 8.9
Mar		3	3	2	8 .2
Apr		101	129	264	494 11.1
May		1		4	5 .1
Jun		99	138	417 ***	654 14.7
Jul		4	9	15	28 .6
Aug		6	6	20	32 .7
Sep		4	7	6	17 .4
Oct	146	157	308	808 ***	1419 31.9
Nov	4	4	9		17 .4
Dec	150	200	323	689	1362 30.6
Column Total	300 6.7	676 15.2	1055 23.7	2416 54.3	4447 100.0

*** LSAT and GRE administrations in the same month.

Each of the 21 row totals in Table 4.1, corresponds to a sample with data reflecting performance on a common set of GRE test questions, without regard to test administration date. The cell entries correspond to 42 subsamples taking a generic form of the GRE and a generic form of the LSAT.

Each of the 37 test occasions noted in Table 4.2, yielded one or more distinct subsets of form-specific observations of performance on GRE items. The number of such subsets corresponds to the number of different generic forms of the GRE taken by sample members tested on particular occasions. Data not shown in the table indicate that for five of the 21 GRE-form subsamples, all GRE data were generated in a single test administration; for the remaining GRE-form subsamples, data are from two or more testing occasions.

Test items from the diverse forms were scored (rights only) using operational scoring procedures; they were then ordered by type and within-section position. The 186 GRE items were divided into 37 item-type parcels. Table 4.3 shows the items included in the parcels, parcel labels, and so on. Note that each parcel includes a mix of items roughly balanced according to ordering within test sections, to parallel the procedure used in developing LSAT parcels, reported earlier (Section 3). The number of items in parcels ranged between four and seven (for two parcels of GRE analytical reasoning items).

Table 4.3.

Description of GRE Item-Type Parcels Used in the Analysis

Parcel	Items In Parcels	Item Type
gsc01	= sc1 sc4 sc7 sc10 sc14	Sentence completion
gsc02	= sc2 sc5 sc8 sc11 sc13	
gsc03	= sc3 sc6 sc9 sc12	
gana01	= ana1 ana5 ana9 ana13 ana18	Analogy
gana02	= ana2 ana6 ana10 ana14 ana17	
gana03	= ana3 ana7 ana11 ana15	
gana04	= ana4 ana8 ana12 ana16	
grc01	= rd1 rd5 rd9 rd13 rd17 rd22	Reading comprehension
grc02	= rd2 rd6 rd10 rd14 rd18 rd21	
grc03	= rd3 rd7 rd11 rd15 rd19	
grc04	= rd4 rd8 rd12 rd16 rd20	
gant01	= ant1 ant5 ant9 ant13 ant17 ant22	Antonym
gant02	= ant2 ant6 ant10 ant14 ant18 ant21	
gant03	= ant3 ant7 ant11 ant15 ant19	
gant04	= ant4 ant8 ant12 ant16 ant20	
gqc01	= qc1 qc7 qc13 qc19 qc25	Quantitative comparison
gqc02	= qc2 qc8 qc14 qc20 qc26	
gqc03	= qc3 qc9 qc15 qc21 qc27	
gqc04	= qc4 qc10 qc16 qc22 qc28	
gqc05	= qc5 qc11 qc17 qc23 qc29	
gqc06	= qc6 qc12 qc18 qc24 qc30	
grm01	= rm1 rm5 rm9 rm13 rm17	Regular mathematics
grm02	= rm2 rm6 rm10 rm14 rm18	
grm03	= rm3 rm7 rm11 rm15 rm19	
grm04	= rm4 rm8 rm12 rm16 rm20	
gdi01	= di1 di3 di5 di7 di9	Data interpretation
gdi02	= di2 di4 di6 di8 di10	
gar01	= gar1 gar7 gar13 gar19 gar25 gar31 gar38	Analytical reasoning
gar02	= gar2 gar8 gar14 gar20 gar26 gar32 gar37	
gar03	= gar3 gar9 gar15 gar21 gar27 gar33	
gar04	= gar4 gar10 gar16 gar22 gar28 gar34	
gar05	= gar5 gar11 gar17 gar23 gar29 gar35	
gar06	= gar6 gar12 gar18 gar24 gar30 gar36	
glr01	= glr1 glr4 glr7 glr10	Logical reasoning
glr02	= glr2 glr5 glr8 glr11	
glr03	= glr3 glr6 glr9 glr12	

Note. Numbers in parcels reflect item position when items from both sections involving a particular item type are considered as a single set (e.g., items1-6 in the second logical reasoning section, become items 7-12 in the consolidated set of 12 logical reasoning items. This pattern of parcel definition can easily be replicated in data from GRE files.

BEST COPY AVAILABLE

For purposes of assessing factors underlying GRE performance in this multi-form sample—and anticipating the need for comparably scaled LSAT and GRE parcels for use in later analyses involving combined parcels from both tests--raw item-type parcel scores from the GRE and raw item-type parcel scores from the LSAT were z-scaled within the 42 subsamples outlined in Table 4.1. For analyses involving the GRE only, reported below, the z-scaled GRE parcels were pooled to yield correlations for analysis.[10]

Keeping in mind the unique nature of the GRE data involved in these analyses—21 different forms of the test, data collected on 37 different testing occasions, and so on—and the apparently novel application of this pooling approach in a factor analytic study involving multiple forms of the GRE (or other similar tests), it is considered particularly noteworthy that, as will be seen below,

(a) factor outcomes involving the full complement of z-scaled GRE item-type parcels, conformed in basic aspects to outcomes that have been reported in previous GRE factor studies—especially a study involving both operational and experimental GRE parcels (Emmerich, Enright, Rock, and Tucker, 1991) in which, as in the present study, item types were not constrained to load on particular factors, and

(b) factor outcomes in LSAT-parallel analyses—involving only the GRE logical reasoning, analytical reasoning, and reading comprehension parcels—generally paralleled the outcomes reported in Section 3 for the corresponding LSAT parcels.

Analytic Procedure

As the initial step in identifying factors underlying performance on the full array of z-scaled GRE parcels, a principal components analysis was made of the corresponding intercorrelations in the total sample (N = 4,447). Results for the first four components are summarized in Table 4.4.

---

[10] Analyses of correlations based on these z-scaled parcels were in all essential respects unmarked by "irregularities"--ill-conditioning, failures to converge, and so on--that if present would raise questions regarding the usefulness of the data set, per se. As a matter of incidental interest, parallel analyses were conducted using the corresponding raw GRE parcels. Again, factor outcomes were free of irregularities. It would appear that form-to-form variation in the properties of particular sets of test items--e.g., level of difficulty, validity--tend to "average out" as the number of forms involved increases. In any event, this incidental finding suggests that a line of inquiry that appears to warrant further research.

Table 4.4.

Summary of Eigenvalues and Associated Percentages of Total Variance

	Principal Component			
	PC1	PC2	PC3	PC4
Eigenvalue	13.06	2.5	1.3	0.9
Percent variance	37.3	7.3	3.9	2.7
Cumulative percent	37.3	4.6	8.5	0.7

The first three components had eigenvalues greater than 1.0, and the fourth eigenvalue was close to 1.0 (0.96), indicating that it would be useful to examine both a three-factor model and a four-factor model. This reasoning is based in part on evidence suggesting that "a practical basis for finding the number of common factors that are necessary, reliable, and meaningful for explanation of correlations among . . . variables . . . is that the number of common factors should be equal to the number of eigenvalues greater than one of the correlation matrix (with unities in the diagonal)" (Harman, 1976: p. 185, citing basic work by Kaiser, 1960). Accordingly, the first three principal components were rotated to orthogonal (varimax) and oblique (direct oblimin) solutions, and the same procedures were used to rotate four principal components factors.

Table 4.5 shows the rotated factor matrix for the three-factor varimax solution, and the corresponding pattern coefficient matrix for the oblimin solution. Note that the ordering of verbal parcels (GANA to GRC) and quantitative parcels (GRM to GDI), respectively, and the placement of the three logical reasoning parcels (GLR), after the verbal parcels, is designed to facilitate overall evaluation. For present purposes in evaluating the three orthogonal factors and the corresponding oblique factors, it is sufficient to note that the two solutions were quite similar with respect to both factor identification and order of emergence of factors.

Table 4.5

Loadings of GRE Parcels on Three Orthogonal Factors, and Corresponding Pattern Coefficients on Three Oblique Factors: Three-Factor Model

GRE parcel	Factor loading Factor 1	Factor 2	Factor 3
GANT01	.65	.19	.06
GANT02	.68	.18	.08
GANT03	.66	.18	.06
GANT04	.62	.17	.08
GANA01	.57	.23	.09
GANA02	.49	.22	.04
GANA03	.51	.21	.10
GANA04	.50	.20	.11
GSC01	.63	.18	.16
GSC02	.60	.17	.18
GSC03	.60	.16	.14
GRC01	.58	.18	.35
GRC02	.56	.18	.38
GRC03	.58	.16	.34
GRC04	.57	.19	.32
GLR01	.48	.15	.34
GLR02	.48	.17	.32
GLR03	.48	.18	.35
GRM01	.22	.69	.24
GRM02	.21	.68	.26
GRM03	.22	.70	.26
GRM04	.22	.69	.26
GQC01	.25	.63	.19
GQC02	.22	.62	.21
GQC03	.26	.66	.19
GQC04	.25	.64	.20
GQC05	.24	.64	.21
GDI01	.25	.53	.33
GDI02	.24	.56	.30
GAR01	.20	.33	.70
GAR02	.20	.33	.68
GAR03	.19	.32	.68
GAR04	.19	.34	.68
GAR05	.20	.32	.68
GAR06	.18	.35	.67

Note. Varimax converged in 5 iterations.

GRE parcel	Pattern matrix Factor 1	Factor 2	Factor 3
GANT01	.70	.04	-.11
GANT02	.73	.01	-.09
GANT03	.72	.02	-.12
GANT04	.67	.01	-.08
GANA01	.58	.11	-.08
GANA02	.50	.13	-.12
GANA03	.52	.10	-.05
GANA04	.51	.08	-.03
GSC01	.67	.00	.00
GSC02	.63	-.00	.04
GSC03	.65	-.01	-.01
GRC01	.60	-.04	.23
GRC02	.57	-.04	.28
GRC03	.61	-.06	.23
GRC04	.58	-.01	.20
GLR01	.48	.04	.26
GLR02	.49	.01	.22
GLR03	.48	.01	.25
GRM01	-.02	.76	.03
GRM02	-.03	.75	.05
GRM03	-.02	.77	.04
GRM04	-.02	.75	.05
GQC01	.05	.69	-.02
GQC02	.02	.68	.01
GQC03	.05	.73	-.03
GQC04	.04	.70	-.01
GQC05	.03	.70	-.00
GDI01	.07	.51	.17
GDI02	.05	.57	.12
GAR01	.05	.17	.66
GAR02	.06	.18	.64
GAR03	.04	.17	.65
GAR04	.04	.19	.64
GAR05	.05	.17	.64
GAR06	.02	.21	.63

Note. Oblimin converged in 12 iterations.

The three factors, in each instance, correspond generally to the three GRE sections: the first factor is strongly identified by the four verbal item-type parcels; the second is clearly identified by the quantitative parcels; and the third is identified primarily by the analytical reasoning parcels.

Of particular interest, for present purposes, is the fact that the three logical reasoning parcels load more heavily on the first factor (defined primarily by verbal item-type parcels) than on the third factor (defined primarily by the analytical reasoning items); but logical reasoning parcels load less strongly on the verbal factor than do the corresponding verbal item types, including reading comprehension.

This particular factor outcome is consistent with previous GRE research findings (especially Schaeffer and Kingston, 1988; Emmerich, Enright, Rock, and Tucker, 1991) and findings from the LSAT analysis reported in Section 3, as well—in suggesting that the analytical reasoning item type and the logical reasoning item type tap somewhat different aspects of reasoning ability. As will be seen below, in the four-factor solution both logical reasoning and reading comprehension shift from the "verbal" factor to define a fourth factor.

Factor loadings for the four varimax factors and pattern coefficients for the four oblimin factors are shown in Table 4.6. The two solutions were similar with respect to order of emergence of factors and factor identification.

Table 4.6.

Loadings of GRE Parcels on Factors: Four-Factor Solutions

GRE parcel (z-scaled)	Factor loading: Varimax factors Factor 1	Factor 2	Factor 3	Factor 4
GANT01	.17	.68	.11	.11
GANT02	.15	.70	.13	.12
GANT03	.15	.68	.10	.12
GANT04	.14	.65	.12	.11
GANA01	.21	.58	.12	.12
GANA02	.21	.50	.06	.12
GANA03	.21	.48	.10	.18
GANA04	.17	.53	.15	.06
GSC01	.18	.56	.12	.31
GSC02	.18	.49	.12	.36
GSC03	.16	.53	.11	.29
GRC01	.19	.45	.27	.43
GRC02	.19	.41	.29	.47
GRC03	.17	.45	.25	.45
GRC04	.20	.45	.25	.41
GLR01	.20	.24	.17	.62
GLR02	.22	.22	.12	.68
GLR03	.22	.24	.17	.63
GQC01	.63	.22	.17	.16
GQC02	.62	.20	.20	.14
GQC03	.66	.24	.18	.14
GQC04	.64	.22	.18	.14
GQC05	.65	.20	.18	.16
GRM01	.68	.21	.24	.10
GRM02	.68	.18	.25	.14
GRM03	.70	.20	.25	.13
GRM04	.69	.20	.25	.13
GDI01	.53	.18	.29	.22
GDI02	.57	.19	.26	.21
GAR01	.32	.17	.70	.19
GAR02	.32	.19	.69	.16
GAR03	.31	.17	.69	.15
GAR04	.32	.18	.70	.14
GAR05	.31	.18	.68	.16
GAR06	.34	.15	.67	.18

Note. Varimax converged in 8 iterations.

GRE parcel (z-scaled)	Pattern coefficients: Oblimin factors Factor 1	Factor 2	Factor 3	Factor 4
GANT01	.01	.73	.00	**-.04**
GANT02	-.03	.76	.03	**-.04**
GANT03	-.01	.74	-.01	**-.02**
GANT04	-.02	.70	.03	**-.03**
GANA01	.08	.60	.01	**.02**
GANT02	.13	.50	-.06	**.00**
GANA03	.10	.46	-.02	**.08**
GANA04	.03	.57	.08	**-.08**
GSC01	.02	.51	-.00	**.25**
GSC02	.03	.41	-.01	**.30**
GSC03	.01	.49	-.01	**.21**
GRC01	-.02	**.32**	**.19**	.37
GRC02	-.02	**.25**	**.21**	.41
GRC03	-.05	**.31**	**.18**	.39
GRC04	.01	**.31**	**.16**	.35
GLR01	.06	-.00	.05	.66
GLR02	.11	-.05	-.03	.74
GLR03	.09	-.01	.04	.67
GRM01	.75	.02	.04	-.07
GRM02	.75	-.03	.05	-.01
GRM03	.77	-.01	.04	-.02
GRM04	.75	-.00	.05	-.02
GQC01	.71	.03	-.04	.02
GQC02	.68	.01	.00	.01
GQC03	.74	.05	-.04	-.00
GQC04	.71	.04	-.03	.01
GQC05	.72	.00	-.03	.04
GDI01	.51	-.02	.15	.11
GDI02	.58	-.02	.10	.09
GAR01	.02	-.02	.78	-.03
GAR02	.02	.02	.78	-.01
GAR03	.01	.00	.79	-.01
GAR04	.03	.01	.79	-.03
GAR05	.01	.00	.77	.00
GAR06	.07	-.04	.75	.02

Note. Oblimin converged in 8 iterations. Underlining indicates the largest pattern coefficient for each parcel. Bolding indicates secondary loadings for the parcels involved on a particular factor.

The ordering of presentation of verbal parcels (GANT, GANA, GSC) and quantitative parcels (GRM, GQC, GDI), and the placement of logical reasoning (GLR) parcels immediately after the reading comprehension (GRC) parcels, is designed to facilitate interpretation. Discussion is focussed primarily on the oblique factors.

1. The first factor to emerge is strongly identified by the quantitative parcels. The magnitudes of the pattern coefficients on this factor have a particular order: highest for the regular mathematics parcels (GRM), next highest for the quantitative comparison parcels (GQC), and lowest for the two data interpretation parcels (GDI).

2. The second factor is identified primarily by the verbal item types. The magnitudes of the pattern coefficients on this factor also have a particular order: highest for antonym parcels (GANT), next highest for analogy parcels (GANA), next highest for sentence completion parcels (GSC), and lowest for reading comprehension parcels (GRC).

3. The third factor is identified primarily by the six analytical reasoning parcels (GAR). Coefficients for reading comprehension (GRC) are relatively high on this factor, as are coefficients for the three sentence completion parcels (SC), and, to a lesser extent, data interpretation (cf. relative magnitudes of the loadings for parcels on the corresponding varimax factor).

4. The fourth factor is identified primarily by the three logical reasoning parcels and the four reading comprehension parcels (GRC). As indicated by their elevated coefficients on the second factor, the reading comprehension (GRC) parcels maintain a verbal identity as well.

Coefficients for the three sentence completion parcels on the fourth factor exhibit tendencies observed in past GRE studies for sentence completion items (which require only sentence-level analysis) to "follow" the more complex reading comprehension item type.

As was true for the first two factors, it can be seen in Table 4.6 that on the fourth factor there is also a particular ordering of the logical reasoning and verbal items types by the magnitudes of the corresponding pattern coefficients: that is, ranked from highest to lowest in terms of the magnitude of the coefficients, we see logical reasoning (GLR), reading comprehension (GRC), sentence completion (GSC), analogy (GANA), and finally antonym (GANT) parcels.

The ordering of the item types on the fourth factor is exactly the reverse of their ordering on the second factor—significant as part of a pattern of similarity between the findings under consideration here and previously reported findings (Emmerich, Enright, Rock, and Tucker, 1991), to be developed in detail later. The fourth factor appears to reflect complex inferential and reasoning skills tapped by reading comprehension items and logical reasoning items, and to a lesser degree by sentence completion items. These item types were associated with the verbal factor in the three-factor solution described earlier.

Among other things, the pattern of outcomes suggests that the fourth factor may reflect a "general reasoning" domain, distinguishable from a "pure verbal reasoning" domain represented by the verbal skills tapped by the discrete-verbal item types—sometimes called "vocabulary" items, which are recognized as relatively pure measures of verbal reasoning. Factors 2 and 4 are also distinguished by the degree to which Factor 4 is defined by item types involving connected discourse.

For interpretive perspective, it is useful to consider in some detail points of similarity, alluded to above, among the four factors described above, and four factors identified by Emmerich, Enright, Rock, and Tucker (1991) in a factor analysis (Promax rotation of principal components) of operational and experimental GRE item-type parcels, in which, as in the present study, " . . . no a priori constraints were placed on the item types that were allowed to define the factors" (p. 53).

Exhibit 4a (Table 17a from Emmerich et al., 1991) provides essential detail regarding loadings of the operational item types (ANT through LR5) and experimental item types (AR3 through CV3) on the four factors identified and evaluated by Emmerich et al.

Exhibit 4a.

Selected Findings of a Study by Emmerich, Enright, Rock, and Tucker (1991): Table 17a from the Study Report

Table 17a

Exploratory Factor Analysis of All Item Types and Parcels
For the 3-Option Multiple-Choice Experimental Sample*

	Factor Loadings**			
General Test	Verbal	Informal Reasoning	Formal-Deductive Reasoning	Quant.
ANT a	.94	-.31	.08	-.03
ANT b	.88	-.21	.09	.06
ANL a	.74	-.04	.02	-.01
ANL b	.72	-.01	-.07	.05
SNCP a	.71	.06	-.07	.11
SNCP b	.62	.08	.01	.10
RCMP a	.54	.24	.16	-.04
RCMP b	.55	.16	.32	-.22
QC a	.07	.04	.46	.38
QC b	-.03	.07	.44	.41
DQ a	.06	-.06	.40	.53
DQ b	-.14	.04	.39	.53
DI a	.04	-.05	-.01	.82
DI b	.02	-.06	.03	.79
AR5 a	.02	-.04	.87	.00
AR5 b	.05	-.08	.90	-.02
AR5 c	.00	.02	.79	.07
LR5 a	.27	.03	.31	.08
LR5 b	.38	.30	-.10	.14
LR5 c	.46	.27	.06	-.12
Experimental Battery				
AR3 a	.06	.09	.71	-.06
AR3 b	.06	.05	.70	.00
LR3 a	.29	.37	.08	.02
LR3 b	.27	.39	.02	.22
NLR3 a	.11	.53	-.09	.32
NLR3 b	.11	.62	-.21	.29
PI3 a	-.24	.75	.33	-.07
PI3 b	-.29	.76	.32	-.03
AX3 a	.24	.63	-.19	-.13
AX3 b	.36	.51	.09	-.28
CV3 a	.57	.18	-.15	.12
CV3 b	.53	.29	.02	-.08

*Principal Components with Promax Rotation.
**Loadings equal to or greater than .30 are underlined.

Precise identification of the experimental item types, for which acronyms are shown in the exhibit, is not essential for present purposes. It is useful to note, however, that LR3, NLR3, AX3 and CV3 were more highly correlated with operational logical reasoning items than with operational analytical reasoning items; CV3 correlated even more highly with the GRE verbal score than with the logical reasoning subscore, and the opposite pattern of relationships obtained for AR3 (a three-option version of the operational AR5-option) item type; PI3 items (number series) were about equally correlated with both logical reasoning and analytical reasoning.

For present purposes, differences in order of emergence of factors are not considered to be pertinent. Attention is directed primarily to similarities in findings, especially,

(a) similarity between the second factor in Table 4.6 and the "verbal" factor in the exhibit, and

(b) similarity between the fourth factor in Table 4.6 and the factor labelled "informal reasoning," in the exhibit.

First, the ordering of the four verbal item types and the logical reasoning item type in terms of pattern coefficients, is similar on the two factors: that is, on the "verbal" factor as on the second factor in Table 4.6, in order from high to low in terms of loadings (coefficients), are analogy, antonym, sentence completion, reading comprehension, and logical reasoning parcels.

Second, the factor labelled "informal reasoning" appears to be similar to the fourth factor in Table 4.6: that is, "informal reasoning" was defined by strong loadings for experimental item type PI3 (involving number series)—known to overlap to about the same extent with analytical reasoning items and logical reasoning items—and experimental item types (for example, NLR3, AX3) known to be relatively highly correlated with existing operational logical reasoning items (LR5 in the exhibit). On this factor, as noted by Emmerich et al., the ordering of the magnitudes of the four verbal item types " . . . is precisely the reverse of that for the verbal factor!" (p. 53). This same reversal outcome is clearly discernible for the fourth and the second factors in Table 4.6.

Emmerich et al. (1991) evaluated the pattern of findings revealed in Exhibit 4a, as follows:

> "The magnitudes of the loadings for these four (verbal) item types had a particular order (on the verbal factor) being highest for ANT, next highest for ANL, next highest for SNCMP, and lowest for RCMP. This ordering of the loadings gives especially heavy weight to the lexical or word-knowledge components of the GRE verbal measure (ANT and ANL), and does so at some expense to the comprehension and inferential components . . . (SNCP and RCMP). At the same time, the loadings for the same four item types on the informal reasoning factor is precisely the reverse of that for the verbal factor! This pattern of outcomes

BEST COPY AVAILABLE

> supports the implication that the _verbal_ and _informal reasoning_ factors uncovered in our analysis represent closely related but distinctive domains" (p. 53).[11]

To extend the apparent points of similarity, the factor labelled "formal-deductive reasoning" in Exhibit 4a, and the third factor in Table 4.6 have in common the important property that they are identified primarily by operational analytical reasoning items. Similarly, the factor labelled "quantitative" in the exhibit and the first factor in Table 4.6 have in common that they are identified by operational quantitative item types. Differences are apparent only with respect to the rank ordering of quantitative item types, in terms of magnitudes of factor loadings on the factors involved (_cf._ coefficients for quantitative item types on Factor 1 and Factor 4 in Table 4.6 with loadings for the "formal-deductive" and "quantitative" factors in the Exhibit).

Finally, it is useful to compare the factor correlations reported by Emmerich et al. with factor correlations generated in the present analysis, shown in Table 4.7. Note that coefficients for the present study (in parentheses) reflect pairings of factors, without regard to order of emergence, based on observed similarities with those from Emmerich et al.

---

[11] No a priori constraints fixing the identification of item types with factors were involved in this analysis. Commenting on this aspect of their analysis, Emmerich et al. observed that " . . . **the absence of such constraints** is one of the reasons why Tables 17a and 17b (factor correlations) seem to us to be especially compelling for purposes of identifying the underlying structure of the reasoning domain" (p. 53, emphasis added).

Table 4.7.

Factor Correlations: Emmerich et al. and Present Study

	Label from Emmerich et al.			
	Verbal	Informal Reasoning	Formal-Deductive	Quantitative
Verbal	1.00	.59	.51	.47
		(.56)	(.46)	(.54)
	f2*	2,4	2,3	1,2
Informal Reasoning		1.00	.60	.51
			(.48)	(.48)
		f4	4,3	4,1
Formal-Deductive Reasoning			1.00	.62
				(.68)
			f3	3,1
Quantitative				1.00
				f1

*Factors in the present study, identified numerically for reporting purposes, for which correlations are shown in parentheses.

The correlation between informal reasoning and formal-deductive reasoning (r = .60) is somewhat higher than that between Factor 4 and Factor 3 in the present study (r43 = .48), but the overall pattern of factor correlations appears to be similar. The correlation between the second and fourth factors in the present study (r24 = .56) is quite similar to that between the verbal and informal reasoning factors (r = .59), and so on.

Generally speaking, given the points of similarity that have been noted, it appears that the structure-related implications suggested above for findings involving factors labelled "verbal" and "informal reasoning" in Exhibit 4a, are also present in findings involving Factors 2 and 4 in Table 4.6. In the present context, apart from their direct substantive implications, findings reported above involving the full array of GRE item types in the selected sample of LSAT/GRE test takers, serve incidentally to establish the viability of the data-set under consideration for purposes of assessing GRE dimensionality—prior to conducting analyses involving only the GRE logical reasoning, analytical reasoning, and reading comprehension parcels designed to parallel the analyses reported for the same three LSAT item types in Section 3.

LSAT-Parallel Analyses

To analyze factors underlying performance on GRE logical reasoning, analytical reasoning, and reading comprehension items, the approach employed paralleled that used in evaluating LSAT dimensionality—that is, an assessment was made of both two-factor and three-factor solutions. Table 4.8 shows salient results of the principal components analysis.

Table 4.8.

Eigenvalues and Associated Percentages of Total Variance: Three Principal Components for GRE Reading Comprehension, Logical Reasoning, and Analytical Reasoning Parcels

	Principal component		
Statistic	PC1	PC2	PC3
Eigenvalue	5.98	1.38	0.81
Percent variance	46.0	10.6	6.2
Cum percent	46.0	56.7	62.9

Orthogonal (varimax) and oblique (direct oblimin) rotations were performed, to two-factor and three-factor solutions, summarized in Table 4.9.

- In the three-factor GRE solution, again as in the corresponding LSAT analysis, analytical reasoning items continue to define a distinct factor, but the coalition between reading comprehension and logical reasoning—evident in the two-factor solution—is abridged by the emergence of separate factors corresponding to these two item types.

Table 4.9.

Two- and Three-Factor Orthogonal (Varimax) and Oblique (Direct Oblimin) Solutions: GRE Reading Comprehension, Logical Reasoning, and Analytical Reasoning Parcels

GRE parcel	Varimax loadings	
	F1	F2
GRC01	.26	.69
GRC02	.28	.69
GRC03	.23	.70
GRC04	.25	.68
GLR01	.22	.63
GLR02	.18	.66
GLR03	.23	.64
GAR01	.76	.27
GAR02	.76	.26
GAR03	.75	.25
GAR04	.76	.26
GAR05	.74	.27
GAR06	.75	.27

Note. Varimax converged in 3 iterations.

GRE parcel	Oblimin pattern	
	F1	F2
GRC01	.72	.03
GRC02	.71	.05
GRC03	.74	-.01
GRC04	.71	.01
GLR01	.67	-.00
GLR02	.72	-.06
GLR03	.68	.01
GAR01	.01	.80
GAR02	-.00	.80
GAR03	-.01	.80
GAR04	-.01	.81
GAR05	.01	.78
GAR06	.01	.79

Note. Oblimin converged in 7 iterations.

Factor correlation

	F1	F2.
F1	1.00	
F2	.60	1.00

GRE parcel	Varimax loadings		
	F1	F2	F3
GRC01	.25	.71	.24
GRC02	.26	.71	.24
GRC03	.22	.73	.23
GRC04	.23	.72	.20
GLR01	.21	.25	.69
GLR02	.18	.24	.76
GLR03	.22	.24	.71
GAR01	.76	.19	.22
GAR02	.75	.22	.16
GAR03	.75	.23	.14
GAR04	.76	.22	.16
GAR05	.73	.23	.16
GAR06	.74	.19	.21

Note. Varimax converged in 5 iterations.

GRE parcel	Oblimin pattern		
	F1	F2	F3
GRC01	.01	.77	.02
GRC02	.03	.76	.02
GRC03	-.03	.80	.01
GRC04	-.01	.81	-.03
GLR01	.02	.02	.74
GLR02	-.03	-.01	.83
GLR03	.03	.00	.77
GAR01	.80	-.04	.06
GAR02	.80	.02	-.01
GAR03	.80	.03	-.05
GAR04	.81	.01	-.02
GAR05	.78	.03	-.02
GAR06	.79	-.04	.05

Note. Oblimin converged in 7 iterations.

Factor correlation

	F1	F2	F3
F1	1.00		
F2	.57	1.00	
F3	.48	.57	1.00

It would appear that in the selected LSAT/GRE sample, the factor structure underlying performance on the GRE logical reasoning, reading comprehension, and analytical reasoning parcels, is very similar to the underlying performance on the LSAT logical reasoning, reading comprehension, and analytical reasoning parcels.

Before combining data from the two tests for exploratory factor analyses, it is useful to examine characteristics of the data set—especially the fact that observations on the LSAT and the GRE are not concurrent—that complicates interpretation of the between-test correlations. Such an examination is provided in the following section.

SECTION 5. BETWEEN-TEST CORRELATIONS IN THE JOINT SAMPLE, AND CONSIDERATIONS INVOLVED IN THEIR EVALUATION

Thus far, attention has focussed exclusively on within-test analyses. In Section 3, LSAT data were analyzed, and the analyses in Section 4 involved only data from the GRE General Test. In the within-test analyses, all correlations analyzed reflected concurrent relationships among the variables involved—scores on item-type parcels, section-level item-type scores, and so on. Correlations between GRE parcels reflected pooled (concurrent) within-form relationships—even though the score data involved, derived from 21 different forms of the GRE, were collected in a total of 37 different GRE administrations between October 1988 and December 1991.

The analyses reported in this section are concerned with questions regarding between-test relationships, reflected in correlations between the reported LSAT scaled scores or LSAT item-type (section) subscores, on the one hand, and GRE scaled scores or GRE item-type subscores, on the other. In these analyses, none of the correlations reflects concurrent relationships between the corresponding LSAT and GRE variables. In some instances, three years separated the LSAT and GRE test administrations, in other instances both tests were taken within less than 10 days.

Table 5.1 shows the distribution of the joint LSAT/GRE sample according to the order in which the two tests were taken and time interval between testing occasions. About one-third of the examinees took the GRE after the LSAT, after intervals ranging from a week to six months. The remainder took the LSAT after the GRE, after intervals ranging from nine days to 36 months; for some 17 percent of the sample, both tests were taken in the same month.

Table 5.1.

Distribution of Sample According to Time Between LSAT and GRE Test Administrations and Ordering of the Two Test Administrations

Order: time between		Code	Frequency	Percent	Cum Percent
GRE after LSAT:	1 - 6 mos.	1	916	20.6	20.6
GRE after LSAT:	7 days	2	592	13.3	33.9
GRE before LSAT:	9 days	3	176	4.0	37.9
	1 - 6 mos.	4	450	10.1	48.0
	7 -12 mos.	5	709	15.9	63.9
	13 -18 mos.	6	447	10.1	74.0
	19 -24 mos.	7	444	10.0	84.0
	25 -30 mos.	8	274	6.2	90.1
	31 -36 mos.	9	439	9.9	100.0
		Total	4447	100.0	100.0

Attention is directed first to a brief, primarily descriptive analysis of differences among subgroups classified by order of test taking and time interval between testing occasions (that is, order/interval subgroups or categories) with respect to both average test performance and between-test correlation.

Next, attention is directed to analyses of between-test correlations involving section-level (item-type) LSAT and GRE subscores, especially scores involving the three item types that are common to both tests. In these analyses, like those involving scaled scores from the two tests, between-test correlations were computed for the order/interval subgroups, such as those shown in Table 5.1. These analyses were concerned, in part, with exploring the possibility that time-related effects may be less pronounced for certain item types (for example, reading comprehension) than for others (for example, analytical reasoning or logical reasoning).

Such a possibility is suggested by findings of studies (for example, Wilson, 1988) involving GRE "repeaters" (candidates who take the GRE more than one time), indicating substantially lower "test-retest" correlations for the GRE analytical scaled score (r's of approximately .75 regardless of time interval) than for either the GRE verbal scaled score or the GRE quantitative scaled score (r's centering around .86 for test-retest intervals ranging between a month and over 15 years). The GRE analytical measure is shorter, and correspondingly somewhat less reliable, than either GRE verbal or GRE quantitative.

In any event, based on evidence of lower test-retest correlations for the GRE analytical measure than for the other two GRE measures, it was considered plausible that similar patterns might be observed for subscores on the corresponding item types.

An incidental objective of the analyses reported in this section is to provide perspective for evaluating outcomes of exploratory factor analyses involving parcels of items from both the LSAT and the GRE, to be reported in Section 6.

Scaled Score Analyses

Differences among the order/interval subgroups with respect to performance on the LSAT and the GRE are evident in Table 5.2. Results of one-way analyses of variance indicate that the mean differences in Table 5.2 are statistically significant. However, the principal contribution to observed differences in every instance is by subgroups taking the GRE after the LSAT, true for all four test variables, and subgroup differences are more pronounced in analyses involving the GRE analytical measure than in analyses involving the other test variables.

- When only subgroups taking the GRE before the LSAT were considered in one-way analyses of variance, the resulting F-ratios were not statistically significant, except for the analysis involving the GRE analytical ability measure (cf. F-ratios shown in Table 5.2 for "total" and "before only"). Findings not shown in the table indicate that the two "GRE after LSAT" subgroups are significantly differentiated, statistically, with respect to performance on the test variables.

- The "GRE after LSAT: 7 days" subgroup registered noticeably higher means on all four test variables than did other subgroups.[12]

---

[12] It is possible that LSAT examinees who take the LSAT and then take the GRE may differ from those who have already taken the GRE when they are tested with the LSAT in terms of basic career orientation. The latter may be thought of as GRE examinees who subsequently take the LSAT, whereas the former may be thought of as LSAT examinees who subsequently take the GRE. Detailed consideration of the characteristics of these two subgroups is outside the scope of this study.

Table 5.2.

Descriptive Statistics for Test-Order/Time-Interval Subgroups:
LSAT and GRE Scaled Scores

Order/interval category	N	LSAT		GRE-V		GRE-Q		GRE-A	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
GRE after LSAT: 1 -6 mos.*	916	153	9.2	560	108	571	123	631	115
GRE after LSAT: 7 days*	592	156	9.4	**581**	**111**	**599**	**127**	**658**	**117**
GRE before LSAT: 9 days	176	155	10.3	562	105	575	136	608	130
GRE before LSAT: 1 - 6 mos.	450	154	9.9	558	115	578	125	611	119
7 -13 mos.	709	155	9.4	551	108	565	126	602	120
13 -18 mos.	447	155	8.8	558	103	565	122	599	118
19 -24 mos.	444	155	9.1	560	109	579	123	604	110
25 -30 mos.	274	154	9.2	536	107	564	123	578	120
31 -36 mos.	439	155	9.3	551	112	572	130	582	121
Total	4,447	155	9.4	559	109	574	126	613	120
F ratio: Total		5.24		5.16		4.05		22.39	
Prob: Total		.0000		.0000		.0001		.0000	
Before only: F ratio*		1.67		1.97		1.14		4.00	
Before only: Prob*		.1238		.0661		.3391		.0005	

* Differences among subgroups taking the GRE after the LSAT only.

The correlations shown in Table 5.3 involve scaled scores on the two tests. The data suggest time-related trends in level of correlation between the respective GRE scaled scores and LSAT scaled score. These time-related trends appear to be somewhat more pronounced for correlations involving the scaled GRE analytical score than for the other correlations—that is, with some exceptions, coefficients tend to be larger for subgroups with shorter between-test intervals than for subgroups with longer intervals, but this tendency appears to be more pronounced for GRE-A than for either GRE-V or GRE-Q. By and large, correlations are stronger in samples taking the respective tests within six months than in the remainder of the sample.

Table 5.3.

Correlation of GRE General Test Scores with LSAT Score, by Order/Time Interval Categories: Scaled Scores for Both Tests

		Correlation with LSAT		
Order/interval	N	GRE-V	GRE-Q	GRE-A
GRE after LSAT: 1 - 6 mos.	916	.78	.71	.75
GRE after LSAT: 7 days	592	.77	.73	.79
GRE before LSAT: 9 days	176	.81	.79	.85
Gre before LSAT: 1 - 6 mos.	450	.80	.70	.80
7 - 12 mos.	709	.77	.72	.73
13 - 18 mos.	447	.73	.71	.74
19 - 24 mos.	444	.75	.65	.75
25 - 30 mos.	274	.72	.62	.72
31 - 36 mos.	439	.73	.69	.78
Total	4,447	.76	.70	.74

Highest correlations obtain for the "GRE-before-LSAT: 9 days" subgroup. As can be seen in Table 5.2, the GRE analytical ability and quantitative ability standard deviations, and the LSAT standard deviation, but not that for GRE verbal, were elevated somewhat in this particular subgroup.

Explication of these trends is outside the scope of the present study. For present purposes, the findings are of interest primarily because they provide evidence of time-related trends in between-test correlations involving scaled scores—trends that appear to be somewhat more pronounced for the GRE analytical measure than for the other two GRE measures. Findings involving LSAT item-type subscores, reported below, indicate that correlations involving reading comprehension subscores tend to be somewhat more stable across order/interval subgroups than correlations involving logical reasoning subscores and, especially, analytical reasoning subscores.

Patterns of Between-Test Correlations Involving Subscores on Item Types Common to Both Tests

Before examining correlational findings involving subscores, a brief overview is provided of the subscores involved and procedures followed in computing and rescaling subscores for the respective tests, and in estimating reliabilities used to generate disattenuated correlations.

Item-Type Subscores Involved in Analyses

The item-type subscores used in the analyses reported in this section are enumerated in Table 5.4. In most instances, the scores involved correspond directly to item-type sections. A brief description is in order for two subscores.

- First, the GRE subscore labelled GRElsat is a composite of GRE logical reasoning, analytical reasoning, and reading comprehension subscores. This GRE subscore reflects performance on the three item types that are common to both the LSAT and the GRE, and thus is LSAT-parallel with respect to item-type composition, but not mix—the GRE subscore is more heavily weighted with analytical reasoning items than with logical reasoning items, while the opposite is true of the LSAT.

- Second, in parallel factor analyses of logical reasoning, reading comprehension, and analytical reasoning parcels, conducted for the LSAT and the GRE, generally parallel two-factor solutions were obtained—one factor in each test analysis was defined by parcels of logical reasoning items and parcels of reading comprehension items. The subscores labelled "informal reasoning" are thought of as specifying these factors *operationally* for each test (although they are not "factor scores," per se).

Table 5.4.

Outline of Item-Type Subscores from the LSAT and the GRE

Item type	GRE	No. items	LSAT	No. items
Logical reasoning	GLR	12	LLR24	24
			LLR25	25
Analytical reasoning	GAR	38	LAR	24
Reading comprehension	GRC	22	LRC	28
GRElsat*	GLR+GAR+GRC	62	(LSAT)	101
"Informal reasoning"**	GLR+GRC "GINFR"	33	LLR+LRC "LINFR"	77
Sentence completion	GSC	14	N.A.	
Analogy	GANA	22	N.A.	
Antonym	GANT	22	N.A.	
Quantitative comparison	GQC	30	N.A.	
Regular mathematics	GRM	20	N.A.	
Data interpretation	GDI	10	N.A.	

* This subscore is a composite of GRE logical reasoning, analytical reasoning, and reading comprehension items—the three item types represented in the LSAT (hence GRElsat).

** This subscore, in both tests, is a composite of logical reasoning and reading comprehension items. These item types define generally parallel factors in LSAT and GRE analyses (reported in Section 3 and Section 4).

As was done in developing scores for item-type parcels, raw number-right scores computed for the variables designated in Table 5.4, were z-scaled within the 42 GRE/LSAT form-samples (see Section 4, Table 4.1, and related discussion).

Raw-score versions of the GRE item-type scores were also analyzed as a matter of incidental interest. Generally speaking, patterns of outcomes involving raw number-right scores on GRE variables, not reported herein, were similar in all essential respects to patterns of outcomes reported in this section for the corresponding z-scaled scores.

Reliability estimates for the LSAT variables are shown in Table 5.5a, and estimates for the GRE variables are shown in Table 5.5b. As noted in the tables, coefficients based on data for the selected sample and published coefficients for different versions of the GRE and the LSAT, respectively, were evaluated. Given the generally comparable nature of the coefficients evaluated—as indicated in the two tables—it was decided to select a typical value for use in corrections for unreliability (the medians shown in the tables, for all coefficients shown in the respective rows).

Table 5.5a.

Estimates of Reliability for LSAT Subscores*

Part Score (Item Type)	Acronym	n Items	Multiple Forms	LSAT Test Analysis		Median Value (All Coefficients)
				June	Oct	
Sample (N) 4,447						
Logical reasoning-1	(LLR24)	24	.78	.78	.77	.78
Logical reasoning-2	(LLR25)	25	.78	.79	.77	.78
Logical reasoning (1+2)	(LLR49)	49	.89	(.87	.88	.88)**
Reading comprehension	(LRC28)	28	.80	.80	.79	.80
Analytical reasoning	(LAR24)	24	.80	.77	.76	.77
LLR+LRC (informal reasoning)	(Linfr)	77	(.91)			(.91)***

* The LSAT section-score estimated reliability coefficients shown under "median value" are medians of Rulon-reliabilities (Rulon, 1939), computed in the selected sample for subsamples taking June and October LSAT forms, respectively (reported under "multiple forms"), and KR-20 reliabilities reported by ETS for the June and October forms of the LSAT in test-analysis samples ("June" and "Oct" columns).

** LLR49 is the sum of LR24 and LR25; reliabilities estimated from test analysis data for the respective sections.

*** "Linformal reasoning" is the simple sum of z-scaled scores for LSAT reading comprehension and LSAT logical reasoning, the two item types defining the first factor in the two-factor LSAT model and the two-factor GRE model: reliability estimated from internally reported data for the two LSAT forms.

Table 5.5b.

Estimates of Reliability for GRE Item-Type Part Scores, Using Multiple-Form Data and Single-Form Data, Respectively

Part Score (Item Type)	Acronym	n Items	Multiple Forms	Single Form Study Sample	Single Form Test Analysis (a)	Single Form Test Analysis (b)	Median Value (All Coefficients)
Sample (N) 4,447							
Sentence Completion	(GSC)	14	.70	.70	.72	.71	.70
Analogy	(GANA)	18	.63	.59	.65	.71	.64
Reading Comprehension	(GRC)	22	.78	.79	.76	.78	.78
Antonym	(GANT)	22	.78	.79	.78	.79	.78
Quantitative Comparison	(GQC)	30	.85	.82	.85	.86	.85
Regular Mathematics	(GRM)	20	.84	.85	.84	.79	.84
Data Interpretation	(GDI)	10	.72	.63	.64	.73	.68
Analytical Reasoning	(GAR)	38	.88	.87	.86	.89	.88
Logical Reasoning	(GLR)	12	.69	.65	.62	.66	.66
GRC+GLR+GAR (LSAT TYPES)	(Glsat)	77	(.90)	Not available			(.90)
GRC+GLR (Ginformal reas)	(Ginfr)	34	(.84)	Not available			(.84)**

* GRE test analysis reliabilities are KR-20 estimates, computed for recent forms of the GRE General Test: Form 3JGR3 (a), 3OGR4 (b). The "multiple forms" Rulon-reliability estimates were based on raw, odd-even scores for the respective item types, without regard to test form (over 21 different forms); the "single form" estimates are based on data for the largest GRE form-sample represented in the study sample (N = 666).

** "Ginformal reasoning" is the simple sum of z-scaled scores for GRE reading comprehension and GRE logical reasoning, the two item types defining the first factor in the two-factor GRE model and the two-factor LSAT model.

Attention is directed first to an evaluation of patterns of observed between-test correlations, and then to an evaluation of disattenuated correlations for subscores involving item types common to both the LSAT and the GRE.

Time-Related Trends in Observed LSAT/GRE Correlations

Given evidence of time-related trends in correlations involving scaled scores on the LSAT and the GRE, correlations involving the subscores shown in Table 5.4 were computed and evaluated across nine order/interval subgroups–that is, the subgroups for which scaled-score correlations were evaluated (see Table 5.3, and related discussion, above).

Correlations are shown in Table 5.6 for selected LSAT and GRE variables. More specifically,

(a) on the LSAT side, the variables involved are the LSAT scaled score, logical reasoning (LLR), reading comprehension (LRC), and analytical reasoning (LAR); note that the LLR coefficients reported in the table are means of the corresponding coefficients for the two LSAT logical reasoning subscores (LLR24 and LLR25);

(b) on the GRE side, the variables involved are the LSAT-parallel subscore (GRElsat), GREinfr (composite of GRE logical reasoning and reading comprehension), GRE logical reasoning (GLR), GRE reading comprehension (GRC) and GRE analytical reasoning (GAR).

Various aspects of the findings reported in Table 5.6 are noteworthy. For example,

(a) correlations tend to peak for subgroups taking both tests after shorter between-test intervals, regardless of order of test-taking,

(b) the LSAT-parallel GRE subscore (GRE lsat) correlates more highly with LSAT scaled score than does any of the GRE scaled scores, indicating the specific underlying affinity between the corresponding LSAT and GRE item types, and

(c) time-related differences in level are more pronounced for correlations involving the LSAT analytical reasoning subscore than for correlations involving the other two LSAT item-type scores.

Table 5.6.

Correlation of LSAT Scaled Score and Item-Type Section Scores with Designated GRE Scores, in Nine Order/Interval Subgroups

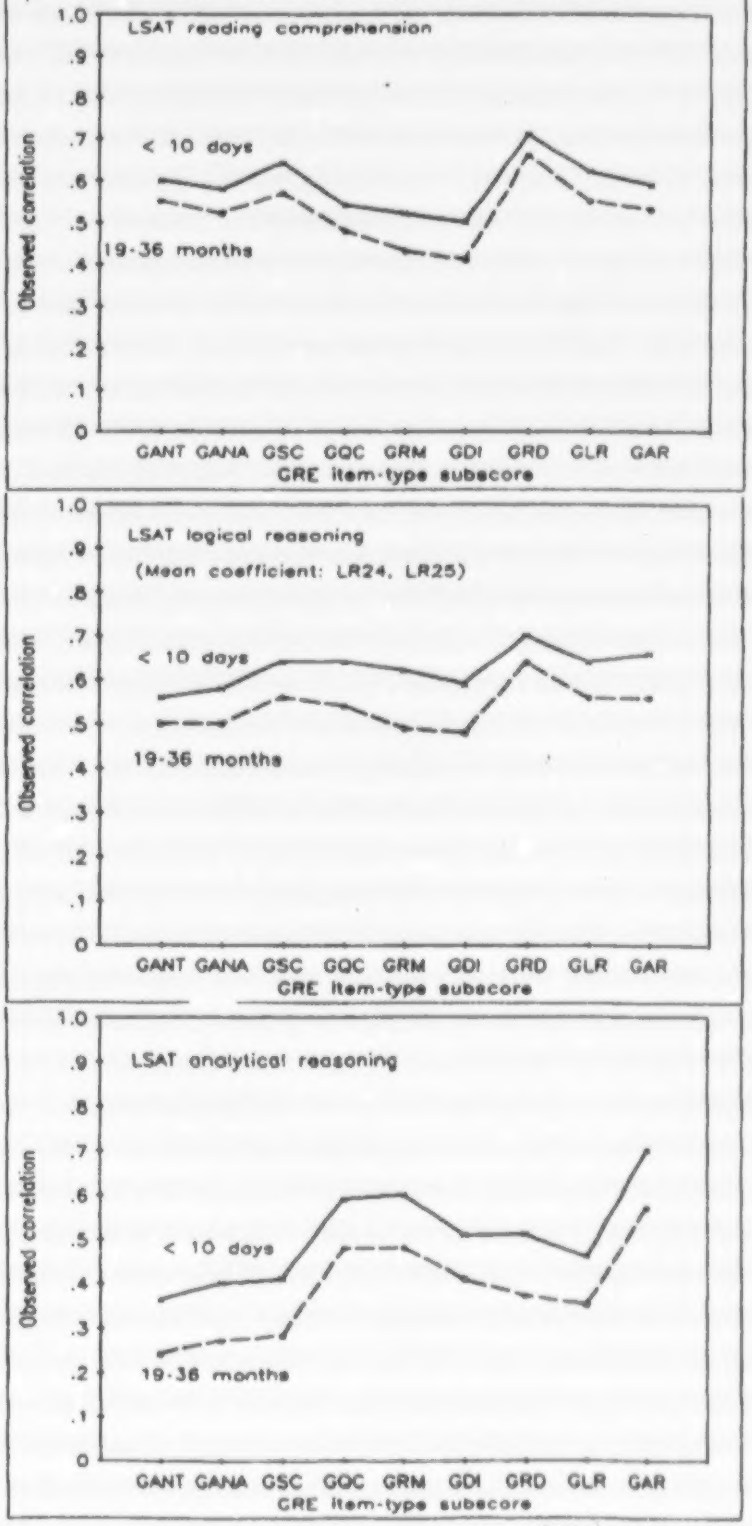
LSAT SCALED SCORE vs.		Designated GRE subscore					GRE scaled score		
	N	(GRELSAT)* GLR+GAR+GRC z-scale	GLR+ GRC	GLR	GRC	GAR	V	A	Q
LSAT was taken . .		z	z	z	z	z			
1 - 6 months after	916	.82	.79	.68	.74	.66	.78	.75	.71
7 days after GRE	592	.84	.79	.68	.75	.74	.77	.79	.73
9 days before	176	.88	.83	.71	.78	.81	.81	.85	.79
1 - 6 months before GRE	450	.82	.78	.66	.76	.74	.80	.80	.70
7 -12 months "	709	.79	.73	.60	.69	.66	.77	.73	.72
13 -18 months "	447	.76	.70	.59	.66	.65	.73	.74	.71
19 -24 months "	444	.80	.74	.61	.74	.69	.75	.75	.65
25- 30 months "	274	.71	.68	.57	.63	.59	.72	.72	.62
31 -36 months "	439	.78	.72	.63	.68	.68	.73	.78	.69
LSAT LOGICAL REASONING (Mean of coefficients for LR24 and LR25)									
1 - 6 months after GRE		.78	.77	.68	.71	.59	.76	.68	.67
7 days "		.82	.78	.67	.75	.69	.77	.75	.71
9 days before GRE		.85	.81	.70	.76	.76	.80	.81	.80
1 - 6 months "		.78	.76	.64	.73	.66	.78	.74	.67
7 -12 months "		.75	.72	.59	.68	.59	.76	.67	.68
13 -18 months "		.71	.68	.56	.64	.57	.71	.65	.66
19 -24 months "		.77	.75	.62	.73	.61	.73	.69	.59
25 -30 months "		.67	.66	.55	.62	.52	.72	.65	.58
31 -36 months "		.78	.74	.66	.70	.66	.74	.76	.68
LSAT READING COMPREHENSION									
1 - 6 months after GRE		.71	.72	.62	.69	.50	.74	.59	.54
7 days "		.74	.73	.61	.71	.57	.74	.64	.57
9 days before GRE		.77	.75	.63	.71	.66	.76	.71	.60
1 - 6 months "		.75	.73	.60	.73	.62	.76	.69	.58
7 -12 months "		.72	.71	.58	.69	.53	.74	.62	.59
13 -18 months "		.65	.64	.52	.61	.49	.70	.59	.53
19 -24 months "		.71	.70	.55	.71	.57	.75	.64	.52
25 -30 months "		.63	.64	.53	.61	.45	.70	.58	.46
31 -36 months "		.70	.67	.57	.66	.55	.70	.65	.54
LSAT ANALYTICAL REASONING									
1 - 6 months after GRE		.64	.52	.45	.50	.67	.48	.68	.64
7 days "		.62	.51	.44	.48	.68	.46	.67	.61
9 days before GRE		.75	.65	.54	.63	.77	.60	.77	.69
1 - 6 months "		.62	.52	.45	.50	.68	.51	.69	.59
7 -12 months "		.57	.45	.38	.41	.62	.45	.63	.59
13 -18 months "		.55	.43	.38	.39	.60	.41	.64	.59
19 -24 months "		.52	.41	.32	.41	.59	.40	.59	.55
25 -30 months "		.52	.41	.39	.34	.57	.36	.62	.54
31 -36 months "		.49	.38	.36	.34	.56	.35	.59	.53

Note. The strongest correlation of a GRE variable with a designated LSAT variable is indicated by an underlined coefficient.

* Simple sum of z-scaled scores for GRE logical reasoning, analytical reasoning, and reading comprehension item types—paralleling the LSAT with respect to item type composition (hence GLSAT).

Salient time-related differentials in the data are captured by contrasting coefficients for two subgroups, namely, a "shorter interval" subgroup (composed of examinees taking both tests after brief intervals of five days and nine days), and a "longer interval" subgroup (composed of examinees who took the GRE 18 to 36 months before taking the LSAT). This is evident in Figure 5.1, which contrasts profiles of correlations, computed in shorter- and longer-interval subgroups, respectively, for LSAT reading comprehension (upper panel), LSAT logical reasoning (middle panel), and LSAT analytical reasoning (lower panel), with the nine GRE item-type subscores. Note on the horizontal axis that GRE reading comprehension (GRC), GRE logical reasoning (GLR), and GRE analytical reasoning (GAR) are placed together on the far right.

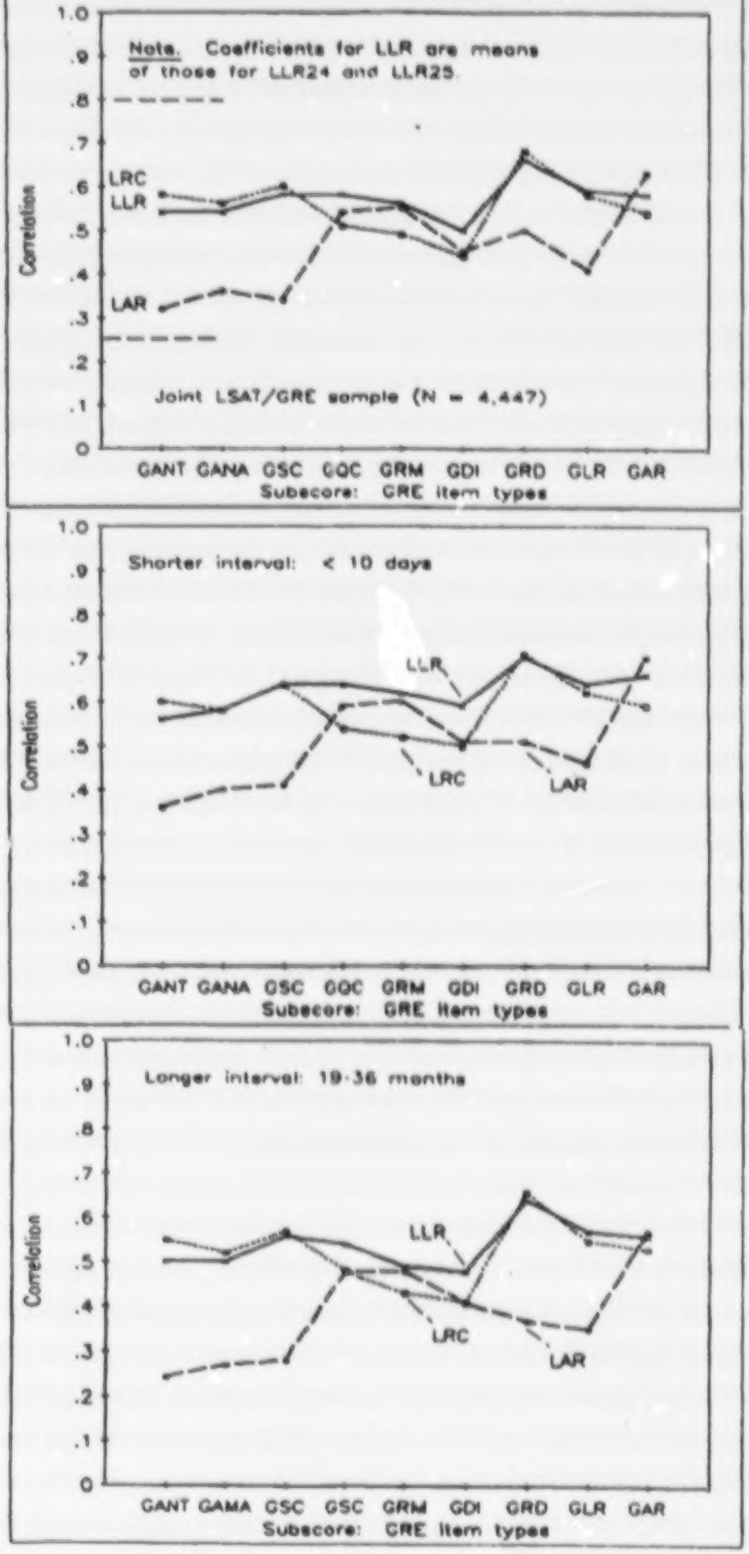
Figure 5.1. Correlation of LSAT reading comprehension, logical reasoning, and analytical reasoning subscores with GRE item-type subscores for subgroups classified by time between tests: < 10 days vs. 19-36 months



For each LSAT variable, the two time-interval correlational profiles are parallel (similar with respect to pattern); they differ only with respect to level. However, it can be discerned that with respect to general levels of correlation, the profiles involving LSAT reading comprehension differ least, the profiles involving LSAT analytical reasoning differ most, while those involving LSAT logical reasoning appear to be in between.

In figure 5.2, correlational profiles for the three LSAT item types are shown together for the total sample (upper panel), the shorter-interval subgroup (middle panel), and the longer-interval subgroup (lower panel). These profiles point up similarities between LSAT reading comprehension (LRC) and LSAT logical reasoning (LLR), especially in correlations involving GRE verbal item types and GRE logical reasoning, as well as clear distinctions between these two item types and LSAT analytical reasoning.

Figure 5.2. Correlation of LSAT logical reasoning (LLR), reading comprehension (LRC), and analytical reasoning (LAR) subscores with GRE item type subscores



LSAT reading comprehension and LSAT logical reasoning appear to diverge a bit, in level of correlations involving the three GRE quantitative item types—that is, LLR correlates more highly with GQC, GRM, and GDI than does LRC, suggesting the possibility that aspects of logical reasoning that are not measured by reading comprehension items, may involve skills that overlap somewhat with skills in the quantitative reasoning domain.[13]

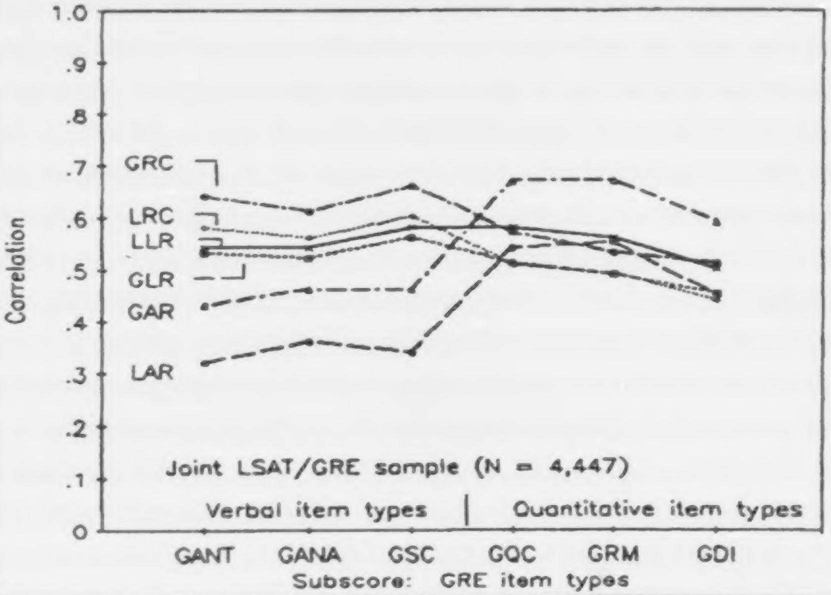
The profiles for LSAT analytical reasoning reflect lower correlations with GRE verbal item types and the GRE logical reasoning subscore than with GRE quantitative item types; and LAR correlates more highly with the analytical reasoning subscore than with any other GRE subscore, a pattern that appears to be least pronounced in the longer-interval subgroup.

To close out consideration of observed correlations, Figure 5.3 shows general parallelism in correlational profiles for LSAT and GRE reading comprehension, logical reasoning, and analytical reasoning subscores, based on correlations between the respective subscores and subscores for the six GRE item types that are not common to both tests: three verbal item types (analogy, antonym, and sentence completion) and three quantitative item types (quantitative comparison, regular mathematics, and data interpretation). In evaluating differences in level of correlation between the respective pairs of item-type subscores, it is important to keep in mind, of course, that the pairs of subscores involved are based on subtests of different length. However, inferences regarding "parallelism" in the correlational profiles are not affected by these differences.

---

[13] The divergence in correlation noted actually appears to be more pronounced for the shorter-interval subgroup than for the longer-interval subgroup--true as well, albeit to a lesser extent, for the correlation of LSAT logical reasoning with GRE analytical reasoning (GAR). This is contrary to the more general tendency for between-test coefficients to be consistently lower in the longer-interval subgroup than in the shorter-interval subgroup. The fact that differences in level of correlations involving LLR and LRC with GRE quantitative items are more pronounced for the shorter-interval subgroup cannot be readily explained. Such differences possibly may be due to differences in relative stability of measurement for the LSAT item types involved. For example, as observed in Figure 5.1, correlations involving LSAT reading comprehension generally are somewhat more stable than are the correlations involving LSAT logical reasoning items. Accordingly, distinctions between logical reasoning and reading comprehension, evident in the shorter-interval profiles, may be "blurred" somewhat in the longer-interval profiles by differential time-related effects: that is, larger relative change in correlations involving logical reasoning than in correlations involving reading comprehension. In such circumstances, only the more proximate observations could be expected to reflect validly any uniqueness in their relationships with other measures. In any event, research is needed to assess the relative stability of abilities measured by these item types.

Figure 5.3. Correlation of LSAT item-type subscores and corresponding GRE item-type subscores with designated GRE verbal and quantitative item-type subscores



The affinity between analytical reasoning items and quantitative items, observed in previous research involving the GRE, is clearly evident for LSAT analytical reasoning items, as well as for the GRE analytical reasoning items. And, the close affinity between reading comprehension and logical reasoning is reflected in their generally parallel correlational profiles.

Estimating Between-Test Overlap

Based on evidence of time-related trends in between-test correlations, as well as logical considerations, data for the shorter-interval subgroup were used in analyses concerned with estimating degree of overlap between LSAT variables and corresponding GRE variables. Corrections for attenuation were made using the standard formula—that is, the observed correlation for each pair of variables was divided by the square root of the product of the median reliability coefficients for the variables involved. Similar corrections were also applied to observed coefficients for the longer-interval subgroup.

Observed and corrected coefficients for designated pairs of LSAT and GRE variables are shown in Table 5.7. Based on the disattenuated coefficients involving the LSAT scaled, there is very substantial overlap between the LSAT and the GRE composite (GRElsat) made up exclusively of GRE logical reasoning, analytical reasoning, and reading comprehension items. The strong disattenuated correlation (r - .93) obtains, even though the proportional mix of the three item types in the GRElsat composite differs substantially from the mix in the LSAT. This holds as well for the two composites labelled "informal reasoning, for which the disattenuated coefficient with GRElsat was also high (r - .93).

Table 5.7.

Summary of Observed and Disattenuated Correlations between Selected LSAT Variables and Designated GRE Variables, for Shorter- and Longer-Interval Subgroups

LSAT VARIABLES vs.	Designated GRE subscore (GRElsat)* GLR+GAR+GRC z-scale z	Ginfr GLR+ GRC z	GLR z	GRC z	GAR z
(Reliability)**	(.90)	(.84)	(.66)	(.78)	(.88)
	Shorter-interval observed correlation				
LSAT (scaled score)	.85 (.92)	.80	.68	.76	.76
LSAT Informal reasoning	.84 (.91)	.81	.69	.78	.70
LSAT Logical reasoning(25)	.76 (.78)	.75	.62	.69	.66
LSAT Logical reasoning(24)	.79 (.78)	.76	.65	.73	.66
LSAT Reading comprehension	.74 (.80)	.73	.62	.71	.59
LSAT Analytical reasoning	.65 (.77)	.54	.46	.51	.70
	Shorter-interval disattenuated correlation				
LSAT (scaled score)	.93	.91	.87	.90	.84
LSAT Informal reasoning	.93	.93	.89	.93	.78
((LSAT Logical reasoning***	.92	.94	.88	.91	.80))
LSAT Logical reasoning(25)	.90	.93	.86	.88	.80
LSAT Logical reasoning(24)	.94	.94	.91	.94	.80
LSAT Reading comprehension	.87	.89	.85	.90	.70
LSAT Analytical reasoning	.78	.67	.65	.66	.85
	Longer-interval observed correlation				
LSAT (scaled score)	.77	.72	.61	.69	.66
LSAT Informal reasoning	.77	.75	.63	.72	.62
LSAT Logical reasoning(25)	.67	.65	.55	.63	.54
LSAT Logical reasoning(24)	.72	.69	.59	.65	.58
LSAT Reading comprehension	.69	.67	.55	.66	.53
LSAT Analytical reasoning	.50	.40	.35	.37	.57
	Longer-interval disattenuated correlation				
LSAT (scaled score)	.85	.82	.78	.81	.73
LSAT Informal reasoning	.85	.86	.81	.86	.69
((LSAT Logical reasoning***	.83	.82	.80	.82	.68))
LSAT Logical reasoning(25)	.80	.80	.77	.81	.65
LSAT Logical reasoning(24)	.86	.85	.82	.83	.70
LSAT Reading comprehension	.81	.82	.76	.84	.63
LSAT Analytical reasoning	.60	.50	.50	.48	.69

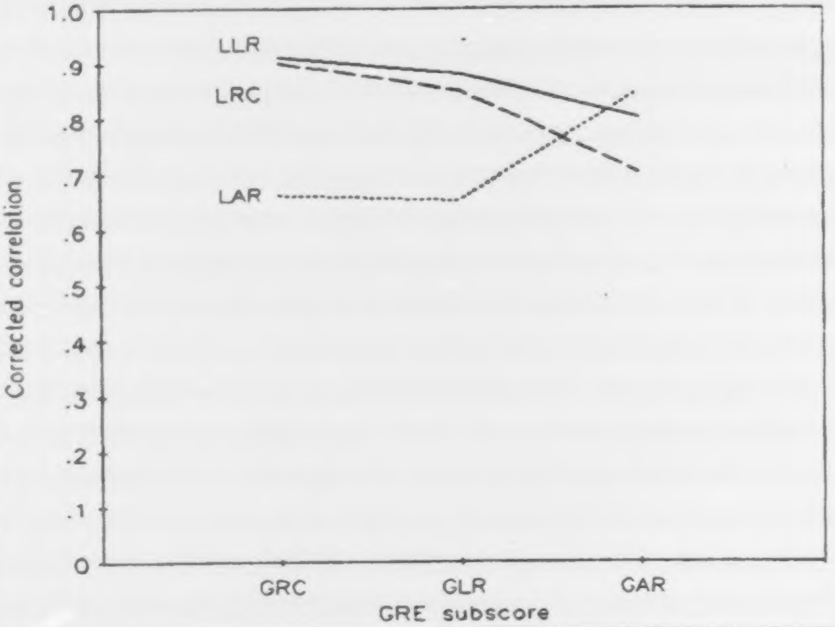
* Simple sum of z-scaled scores for GRE logical reasoning, analytical reasoning, and reading comprehension item types—paralleling the LSAT with respect to item type composition (hence GRElsat).

** See Table 5.5a and 5.5b for detail.

*** This is the rounded mean of corrected coefficients for LLR24 and LLR25.

Figure 5.4 points up the levels and patterns of disattenuated relationships between LSAT item-type subscores and the corresponding GRE item-type subscores in the shorter-interval subgroup. Note that the rounded mean of disattenuated coefficients reported for the two LSAT logical reasoning sections is used in the figure. For LSAT reading comprehension (LRC) and logical reasoning (LLR), the pattern of estimated coefficients involving the three GRE item types is quite similar; and the pattern of estimated coefficients for LSAT analytical reasoning (LAR) is noticeably different—evidencing the psychometric distinctiveness of this item type. Differences in level need to be evaluated in light of evidence reviewed earlier in this section, indicating a particular ordering of the three item types with respect to degree of stability of their correlations with other variables across order/interval categories, namely, from more to less stable, reading comprehension, logical reasoning, and analytical reasoning.

Figure 5.4. Corrected correlations between LSAT and GRE subscores for item types common to both tests



BEST COPY AVAILABLE

Thus, differences in relative levels of correlation involving LRC, LLR, and LAR with their GRE counterparts, indicated in Figure 5.4, appear to be less noteworthy than is the overall pattern of outcomes reviewed in this section, and captured in the figure—a pattern suggesting that

(a) LSAT logical reasoning and reading comprehension item types and their GRE counterparts measure, in common, closely related aspects of general (informal) reasoning ability, and

(b) in both the LSAT and the GRE, the ability domain tapped by logical reasoning and reading comprehension item types is psychometrically distinguishable from that tapped in common by LSAT analytical reasoning items and their GRE counterparts.

More generally, considering the fact that the two sets of test observations involved were not collected on the same testing occasion, it appears likely that the disattenuated coefficients reported in Table 5.7 understate the extent to which the logical reasoning, analytical reasoning, and reading comprehension item types included in the LSAT and their counterparts in the GRE General Test are functionally equivalent in their measurement properties.

The time-related patterns of between-test correlational findings reported in this section for LSAT and GRE scaled scores and section-level item-type subscores from the two tests provide perspective for evaluating findings, reported in the next section, of a factor analysis involving both LSAT item-type parcels (analyzed separately, as reported in Section 3) and the corresponding GRE item-type parcels (also analyzed separately, as reported in Section 4).

SECTION 6. FURTHER EVIDENCE OF COMMON STRUCTURE IN ITEM TYPES COMMON TO THE LSAT AND THE GRE

Findings involving between-test correlations for LSAT and GRE scaled scores and especially those involving section-level item-type subscores from the two tests, reported in the preceding section, are consistent with findings of separate within-test factor analyses, reported in Section 3 and Section 4, in suggesting that

(a) LSAT logical reasoning and reading comprehension item types and their GRE counterparts measure closely related aspects of "reasoning ability" (general or informal, in nature), and

(b) in both the LSAT and the GRE, the abilities tapped in common by logical reasoning and reading comprehension item types are psychometrically distinguishable from those tapped, by LSAT analytical reasoning items and their GRE counterparts.

In within-test factor analyses involving LSAT item-type parcels and the corresponding GRE item-type parcels, two-factor and three-factor models were generally parallel. If the LSAT and GRE observations here under consideration had been obtained in a single testing session, it would follow that two-factor and three-factor models, paralleling those identified in the within-test analyses with respect to item-type specification, should also be identified when the two sets of item-type parcels involved are treated jointly.

However, as we know, the two sets of observations were obtained in different testing sessions. And, as shown in the preceding section, although profiles of between-test correlations involving LSAT item-type section-scores and profiles involving corresponding GRE item-type section scores were generally similar with respect to pattern, they differed systematically in level due largely, perhaps, to influences stemming from the fact that the two sets of test observations were not concurrent, rather than from differences in the respective item types.

Thus, when item-type parcels from both tests are combined for analysis, as in the present section, effects associated with the systematic attenuation of between-test correlations relative to within-test correlations, involving particular item types, can be expected to affect factor outcomes.

Indeed, such effects are revealed clearly in the findings of the exploratory analysis of factors underlying performance on combined parcels of items of the three types that are common to both the LSAT and the GRE General Test, reported in this section.

As will be seen, the findings reflect both

(a) the fundamental similarities already established, with respect to both the within-test and the between-test correlational patterns for the three item types, and

(b) secondary correlational effects attributable to the use of non-concurrent test observations—and associated, systematic between-test differences in levels of correlation involving the same item types.

Analytical Rationale and Procedure

A total of 33 item-type parcels from the LSAT and the GRE, combined—described in detail in Section 3 and Section 4, and outlined briefly below—was available for the combined analysis.

Item Type and Number Items	Parcel Acronym						Number Parcels
Logical reasoning							
LSAT (LLR) 25	LLR51	LLR52	LLR53	LLR54	LLR55		5
LSAT (LLR) 24	LLR61	LLR62	LLR63	LLR64	LLR65		5
GRE (GLR) 12	GLR01	GLR02	GLR03				3
Reading comprehension							
LSAT (LRC) 28	LRC1	LRC2	LRC3	LRC4	LRC5		5
GRE (GRC) 22	GRC01	GRC02	GRC03	GRC04			4
Analytical reasoning							
LSAT (LAR) 24	LAR1	LAR2	LAR3	LAR4	LAR5		5
GRE (GAR) 36	GAR01	GAR02	GAR03	GAR04	GAR05	GAR06	6

Intercorrelations of z-scaled scores on the 33 parcels were computed. A principal components analysis was conducted, and the first six components were retained for exploratory analysis.

The decision to retain six components for further exploratory analysis was made in order to permit a direct empirical assessment of the extent to which factor structure in analyses involving combined parcels from both tests reflects both

(a) the primary underlying relationships—suggesting common underlying structure—already evidenced by strong similarities in patterns of between-test correlation involving the three item types as well as by the parallel outcomes of within-test factor analyses, and

(b) secondary correlational trends associated with systematic differences in levels of between-test correlations involving item-type section scores, that would lead to factorial distinctions between corresponding pairs of item types from the respective tests—distinctions that are largely nonconstruct-related.

It is useful to recall, in connection with the foregoing, that in the separate analyses of the two sets of parcels, three principal components were retained to permit evaluation of both two-factor and three-factor models—of which the latter represents the ultimate decompositional resolution of intercorrelations involving only three types of items within a given test.

Given the between-test influences that have been identified in the current dataset, attainment of the related objectives noted above calls for extracting, rotating, and evaluating six factors, thus permitting ultimate decomposition of the correlation matrix into factors corresponding to the six test/item-type parcel sets reviewed briefly above. It is important to keep in mind, of course, that the extension of factorization in this way is designed primarily to permit assessment of the mix of within-test and between-test correlational effects that is peculiar to this particular data set.

Table 6.1 documents loadings of parcels on the first six principal components extracted using data for the total sample, and the associated eigenvalues. Three of the latter were greater than 1.0 and two were close—eigenvalues of .9 for components four and five. Recalling that only two eigenvalues were equal to or greater than 1.0 in separate within-test analyses involving these parcels, the results in Table 6.1, alone, indicate somewhat greater factorial complexity for intercorrelations involving the combined parcels—due to the mix of within-test and between-test correlational effects peculiar to this sample—than was evidenced in the separate within-test analyses.

Table 6.1.

Summary of Selected Results of Principal Components Analysis: Eigenvalues and Associated Percentage of Total Variance: Six Components

	Principal Components					
Statistic	PC1	PC2	PC3	PC4	PC5	PC6
Eigenvalue	12.69	2.14	1.37	.91	.90	.77
Percent var	38.5	6.5	4.2	2.7	2.7	2.3
Cum percent	38.5	45.0	49.1	51.9	54.6	56.9

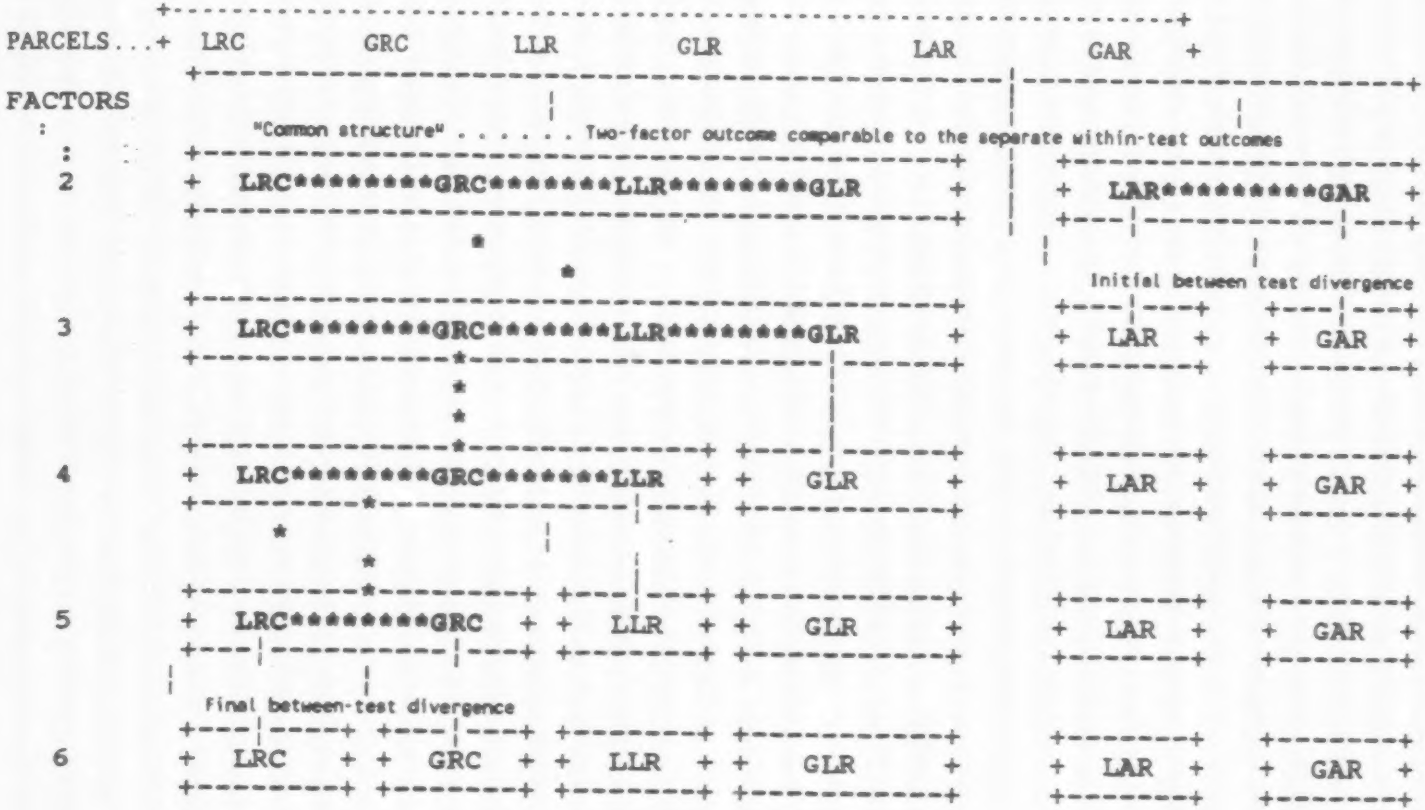
In any event, consistent with the objectives noted above, successive orthogonal (varimax) and oblique (direct oblimin) rotations of corresponding principal components (from Table 6.1), were computed to achieve solutions involving two to six factors. The varimax and oblimin solutions differed only with respect to order of emergence of factors, beginning with the four-factor solution. Only the oblimin factors are discussed in detail. As a matter of incidental interest, the same set of procedures was followed in conducting parallel analyses of intercorrelations of these parcels in the "closer interval" and "longer-interval" (between-tests) subgroups (less than 10 days vs. 19–36 months). Due to general similarities between outcomes of analyses in the two time-interval subgroups and analyses involving the total sample, except as otherwise indicated, only the total-sample findings are considered in detail.

Findings

Basic outcomes of the two-, three-, four-, five-, and six-factor solutions involving data for the total sample, are summarized in Exhibit 6a. LSAT logical reasoning, reading comprehension and analytical reasoning parcels are denoted by LLR, LRC, and LAR, respectively, and the corresponding GRE parcels are denoted by GLR, GRC, and GAR.

Exhibit 6a.

General Pattern of Factorial Decomposition of Intercorrelations of LSAT and GRE Parcels as Indicated by Item-Types Defining Factors Identified in Two-, Three-, Four-, Five, and Six-Factor Oblique Solutions


PARCELS...
LRC
GRC
LLR
GLR
LAR
GAR
FACTORS
"Common structure" . . . . . . Two-factor outcome comparable to the seperate within-test outcomes
2
LRC********GRC*******LLR********GLR
LAR*********GAR
Initial between test divergence
3
LRC********GRC*******LLR********GLR
LAR
GAR
4
LRC********GRC*******LLR
GLR
LAR
GAR
5
LRC********GRC
LLR
GLR
LAR
GAR
Final between-test divergence
6
LRC
GRC
LLR
GLR
LAR
GAR

In the exhibit, patterns of coalescence and divergence of item-type parcels, within tests as well as between tests, can be traced. Parcels coalescing to identify factors in the respective solutions are joined by asterisks and are in bold-face type. In the exhibit, divergence of a parcel is indicated by vertical ("|") linkage with its location in the preceding factor, and separate "boxing."

It can be seen that with each successive analysis, the additional factor extracted is defined by a particular item type, which has diverged (split away) from the type(s) with which it had coalesced to form the preceding factor.

Note that the initial two-factor solution involving combined GRE and LSAT parcels, appears to be directly comparable—based on factor identification by item types—to the respective within-test, two-factor solutions reported earlier (in Section 3, for LSAT item-type parcels and the corresponding three GRE item-type parcels, in Section 4).

The initial between-test divergence involves the two sets of analytical reasoning parcels, LAR and GAR. With the extraction of three factors, LAR and GAR split from their alliance in the two-factor solution to form separate factors, but the between-test coalescence of reading comprehension and logical reasoning items is sustained.

In the four-factor solution, one of the two logical reasoning sets (GLR) diverges to form a separate factor; the second logical reasoning set (LLR) diverges from a sustained alliance with reading comprehension to form the fifth factor. Finally, the longest-sustained association—that involving LSAT and GRE reading comprehension items—is disrupted as these two item types finally diverge to identify separate factors in the six-factor model. The scenario outlined above applies generally to outcomes of analyses involving the shorter-interval and longer-interval subgroups, in which the primary differences had to do with differences in "order of divergence" for particular item types in the four- and five-factor solutions.

In evaluating the sequence of divergence outlined in the exhibit, note that the order of between-test divergence of the item types—first analytical reasoning, next logical reasoning, and then reading comprehension—as additional factors are extracted is consistent with expectation based on the differences in between-test correlations as a function of time between testing occasions (from Section 5). That is, order of between-test divergence for the item types is exactly the reverse of their ordering with respect to resistance to time-related attenuation.

More specifically, we saw in Section 5 that profiles of between-test correlations involving reading comprehension subscores differed least in level, and the exhibit shows that parcels of this item type retained their factorial allegiance longer than did parcels of the other two types. On the other hand, after demonstrating their basic underlying affinity by coalescing to form a single factor in the two-factor model, the factorial alliance between GRE and LSAT analytical parcels (GAR and LAR) was terminated with the extraction of the third factor, consistent with the fact that differences in level of between-test correlation were more pronounced for these item types than were corresponding differences involving the other two item types.

It is of incidental interest too that in the five-factor solution involving the combined parcels, distinctions between reading comprehension item types and logical reasoning types, as well as between these item types and analytical reasoning emerge, short of the final decomposition into factors uniquely defined by the six test/item-type sets. In the separate within-test analyses, these distinctions were captured in three-factor models.

Evidence of Common Structure

Although results of the extended factorization outlined above clearly reflect the between-test correlational effects that were identified and evaluated in Section 5, and are pertinent to overall evaluation of the data under consideration, results of the two-factor solution are more directly pertinent to the common-structure issue.

- In the two-factor solution, basic underlying structural relationships among the three item types emerge, despite the fact that the two sets of test observations are not concurrent—that is, the two-factor solution involving combined LSAT and GRE parcels, tends generally to parallel the separate within-test solutions.

The parallel nature of the findings alluded to is clearly discernible in Table 6.2, which provides detail regarding pattern coefficients for LSAT and GRE item-type parcels on two oblique factors generated (a) in the combined-parcels analysis and (b) in the two separate within-test analyses. Results for the LSAT (first panel) are based on the general LSAT sample (N = 7,385), results for the separate analysis of GRE parcels are shown in the second panel, and the third panel shows results when LSAT and GRE parcels were combined for analysis (cf. summary of the corresponding two-factor solution in Exhibit 6a). The case for "common structure" appears to be a strong one.

Table 6.2.
Pattern Coefficients for LSAT and GRE Parcels on Oblique (Correlated) Factors in Separate Analyses and in Analyses Involving Combined Parcels from Both Tests

Separate LSAT analysis
(General sample, N = 7,385)

LSAT parcel	Pattern F1	Pattern F2
LLR51	.61	.05
LLR52	.63	.06
LLR53	.62	.01
LLR54	.63	.07
LLR55	.62	.05
LLR61	.64	.03
LLR62	.62	.06
LLR63	.66	.06
LLR64	.64	.06
LLR65	.54	.10
LRC1	.72	-.05
LRC2	.69	-.08
LRC3	.72	-.06
LRC4	.67	-.07
LRC5	.68	-.08
LAR1	.04	.72
LAR2	.01	.72
LAR3	-.01	.74
LAR4	.05	.75
LAR5	-.02	.76

LAW SCHOOL ADMISSION TEST

Separate GRE analysis
(Selected LSAT/GRE sample, N = 4,447)

GRE parcel	Pattern F1	Pattern F2
GLR01	.67	-.00
GLR02	.72	-.06
GLR03	.68	.01
GRC01	.72	.03
GRC02	.71	.05
GRC03	.74	-.01
GRC04	.71	.01
GAR01	.01	.80
GAR02	-.00	.80
GAR03	-.01	.80
GAR04	-.01	.81
GAR05	.01	.78
GAR06	.01	.79

GRE GENERAL TEST

Combined LSAT/GRE analysis
(Selected LSAT/GRE sample, N = 4,447)

Parcel	Pattern matrix F1	Pattern matrix F2	
LLR51	.56	.10	LSAT
LLR52	.60	.10	•
LLR53	.62	.00	•
LLR54	.66	.02	•
LLR55	.61	.04	•
LLR61	.63	.03	•
LLR62	.64	.05	•
LLR63	.66	.05	•
LLR64	.62	.07	•
LLR65	.55	.07	LSAT
GLR01	.57	.02	GRE
GLR02	.58	.01	•
GLR03	.58	.04	GRE
LRC1	.72	-.04	LSAT
LRC2	.72	-.07	•
LRC3	.74	-.08	•
LRC4	.69	-.07	•
LRC5	.71	-.08	•
GRC01	.66	.02	GRE
GRC02	.67	.02	•
GRC03	.68	-.02	•
GRC04	.66	-.01	GRE
			¦
LAR1	.02	.66	LSAT
LAR2	-.04	.67	•
LAR3	-.05	.69	•
LAR4	-.01	.71	•
LAR5	-.07	.71	•
GAR01	.05	.73	GRE
GAR02	.08	.68	•
GAR03	.07	.67	•
GAR04	.06	.70	•
GAR05	.08	.68	•
GAR06	.04	.72	GRE

COMBINED LSAT/GRE ANALYSIS

In essence, the study findings suggest common structure for logical reasoning, reading comprehension, and analytical reasoning item types, regardless of the test in which they are used. The structure appears to involve two dimensions, one represented by the item types that measure general reasoning skills that appear to be involved in the analysis of extended discourse (logical reasoning and reading comprehension), and the other represented by more narrowly constrained, formal deductive aspects of reasoning such as those measured by the analytical reasoning item type.

Study findings are summarized briefly and some of their implications are discussed in Section 7.

SECTION 7. STUDY FINDINGS AND IMPLICATIONS: REVIEW AND EVALUATION

The present study was undertaken to help clarify the internal structure of the LSAT, and shed light on the classification of the ability or abilities measured by LSAT reading comprehension, logical reasoning, and analytical reasoning item types, using data from two regularly scheduled LSAT administrations, namely, the June 1991, and October 1991 administrations.

For broader perspective, the study employed data for a subsample from the same two LSAT administrations, composed of LSAT examinees identified as having also taken the GRE General Test between October 1988 and December 1991. Within this sample it was possible to conduct parallel within-test analyses involving item types common to both tests, examine between-test correlations involving scaled scores and specially computed item-type subscores, and ultimately to identify factors underlying performance on parcels of items of three types that are common to both tests, using combined LSAT and GRE data.

The study also drew on the substantial body of evidence regarding relationships among these item types based on research in the GRE context involving GRE reading comprehension, logical reasoning, and analytical reasoning items similar to those used in the LSAT. These three item types have been included in all editions of the GRE since October 1981.

To attain study objectives, a series of related analyses reported in detail in the preceding sections, was undertaken. This section provides a brief recapitulation of these analyses and related findings and considers several implications of the findings.

Recapitulation

First, separate analyses were made of within-test correlations involving primarily scores on item-type parcels in

(a) LSAT data for a general sample of LSAT examinees and data for a selected sample composed of LSAT examinees who also took the GRE General Test, and

(b) GRE data for the selected LSAT/GRE sample.

Findings of separate exploratory within-test factor analyses based on parcels of items of the three types common to both tests were generally parallel. In each analysis, when two factors were extracted, logical reasoning and reading comprehension items coalesced to form one factor, while the second factor was identified primarily by analytical reasoning parcels.

The findings of the factor analyses suggested that these item types in both tests measure psychometrically distinguishable aspects of reasoning ability: aspects of general or informal reasoning, defined by reading comprehension items and logical reasoning items, on the one hand, and aspects of formal, deductive reasoning, defined by analytical reasoning items, on the other.[14]

Next, analyses were made of between-test correlations involving reported, scaled scores and specially computed item-type section scores. These analyses were designed in part to assess effects associated with the fact that the LSAT observations and the GRE observations were collected on different testing occasions separated by intervals ranging from less than 10 days to 36 months. The between-test analyses included assessment of time-related effects on between-test correlations involving the three item types common to both tests.

In these analyses, profiles of correlations involving LSAT item types and their GRE counterparts, computed for shorter-interval (between tests) and longer-interval subgroups (< 10 days versus 19–36 months) were found to be strikingly similar with respect to pattern. They differed with respect to level, reflecting attenuating effects due largely, perhaps, to influences associated with the fact that data for the two tests were not collected concurrently and that time between testing occasions varied substantially. Time-related differences in levels of between-test correlations involving section-scores on the three item types, were smaller for reading comprehension items than for logical reasoning items and, especially, analytical reasoning items.

Results of the analysis of between-test correlations—observed and corrected for attenuation due to the presence of measurement error—involving item types common to both tests, were consistent with the findings of the separate within-test factor analyses, in suggesting psychometrically distinguishable differences between aspects of general or informal reasoning measured by reading comprehension and logical reasoning item types, on the one hand, and aspects of formal, deductive reasoning tapped by the analytical reasoning items, on the other.

---

[14] A similar conclusion was reached by Camilli, Wang, and Fesq (1992), in a study involving data for six forms of the LSAT administered between June 1989 and October 1990. Consistency in findings regarding LSAT dimensionality in studies involving different samples and different methodological approaches, strengthens the conclusion that LSAT analytical reasoning items tap aspects of reasoning ability that are psychometrically distinguishable from those tapped by LSAT logical reasoning and reading comprehension items.

Finally, intercorrelations involving the combined set of LSAT and GRE logical reasoning, reading comprehension, and analytical reasoning item-type parcels from the separate within-test analyses were analyzed. Two-, three-, four-, five-, and six-factor models were computed and evaluated. Extended factorization was needed to permit direct assessment of both

(a) primary underlying structural relationships, evidenced by strong similarities in patterns of between-test correlations involving the three item types as well as by the parallel outcomes of within-test factor analyses, and

(b) secondary correlational trends associated with systematic differences in levels of between-test correlations involving item-type section scores from the two tests, that by inference can be expected to result in factorial distinctions between corresponding item-types from the respective tests with the extraction of additional factors—distinctions reflecting time-related between-test influences that are perhaps largely nonconstruct-related.

Results of the extended factorization clearly illuminated the between-test influences. For example, between-test divergence of particular item-type parcels (splitting into separate factors, after having coalesced to define a preceding factor) occurred first for LSAT and GRE analytical reasoning items, next for the two sets of logical reasoning parcels, and last for the two sets of reading comprehension parcels.

This particular ordering of the item types with respect to timing of divergence, is exactly the inverse of their ordering in terms of resistance to time-related attenuation: that is, analytical reasoning items were least resistant, reading comprehension items were most resistant, and logical reasoning items were in between.

Although results of the extended factorization concerned with between-test effects are pertinent to an overall evaluation of the data, results of the two-factor solution involving the combined parcels are most directly illuminating for study purposes.

In the two-factor solution, basic underlying structural relationships among the three item types emerged, despite the fact that the two sets of test observations are not concurrent—that is, the two-factor solution involving combined GRE and LSAT parcels generally paralleled the corresponding within-test solutions.

In essence, the study findings suggest common structure for LSAT and GRE logical reasoning, reading comprehension, and analytical reasoning item types, regardless of the test in which they are used. The structure appears to involve two dimensions, one represented by the item types that measure general reasoning skills that appear to be involved in the analysis of extended discourse (logical reasoning and reading comprehension), and the other represented by more narrowly constrained, formal aspects of reasoning such as those measured by the analytical reasoning item type.

As noted at the outset, if a test (such as the LSAT) employs more than one type of item and/or different kinds of content for a given item type, the extent to which the different types of test questions tap different aspects of a particular ability, or in fact tap different aspects of underlying abilities, needs to be assessed. The findings that have been reviewed represent the outcomes of one such assessment. The findings have implications for both testing programs.

Implications

For the LSAT as presently constituted, perhaps the most central conclusion supported by the findings is that

- the logical reasoning, reading comprehension, and analytical reasoning item types included in the LSAT, have potential to generate more information than is now being conveyed by the single LSAT scaled score.

That potential is suggested by findings indicating that (a) the LSAT item types measure psychometrically distinguishable aspects of reasoning ability, and (b) the attendant possibility that the information provided by item-type subscores might prove to be useful for predictive or diagnostic purposes in the law school admission context[15]—for example, a score based on logical reasoning and reading comprehension items and a score based on the analytical reasoning items, consistent with the basic two-factor outcomes.

Of immediate interest are questions concerning differential validity of scores based on item types for predicting first-year law school grades generally, grades in particular courses or clusters of courses—for example, clusters defined by a priori judgment as to differences in demands on types of reasoning skills tapped by LSAT items, grades in successive years of legal education, and so on.

---

[15] Such a possibility, appears to have been anticipated by Wightman and Muller (1990b), in their study of the comparative validity of the LSAT scaled score (on a pre-June 1989 version of the LSAT which included the three item types in the current version) for members of different ethnic groups. This study included an evaluation of the means of ethnic groups on item-type sections as well as on the LSAT scaled score. "The data do suggest important areas of further inquiry that should be pursued. For example, minority students continue to perform more poorly than white students not only on the predictors, but also on the criterion variable, first-year average. The predictor variables, especially the LSAT score need to be evaluated carefully to try to determine whether important diagnostic information can be extracted from the scores" (pp. 27-29, emphasis added). In this study, ethnic group differences in patterns of means on item-type subscores were evident. The comparative validity of the respective subscores for predicting first-year law school grades was not at issue in the study.

A study of the comparative validity of LSAT section scores for predicting such criteria in general samples, and in samples defined by ethnic group membership, age,[16] gender, undergraduate major, and so on, would contribute toward resolution of academically, psychometrically, and socially important "differential validity" questions in the current LSAT context.

Closely related to the foregoing are questions concerning incremental validity, especially as to whether differential weighting of item-type subscores results in better prediction of pertinent criteria (for example, grades in successive years of legal education), than is provided by the total scaled score which is based on the simple sum of correct responses to test items—in general samples of students and in subgroups defined by ethnicity, age, undergraduate major field, and/or gender.[17]

Study findings indicate that between-test correlations involving reading comprehension subscores are less attenuated by time-related influences than are subscores for logical reasoning and analytical reasoning—the latter being least resistant to time-related attenuating effects. Such findings are consistent with findings in the GRE context (for example, Wilson, 1988) indicating that "stability coefficients" (test-retest correlations for GRE repeaters) for the GRE analytical measure are some .10 correlation points lower than those for either the GRE verbal measure or the GRE quantitative measure (test-retest correlations centered around .75 for the GRE analytical ability versus .86 for the other two measures).

The GRE analytical measure is shorter, hence somewhat less reliable than the other two measures. However, in the present study, the time-related differences alluded to above emerged in both observed correlations and correlations corrected for attenuation due to measurement error; and in the analyses involved, logical reasoning and analytical reasoning subscores were treated separately, rather than jointly, as in the GRE analytical ability measure.

Based on findings of the present study—in which inferences about "relative stability" necessarily were based on between-test correlations involving comparable item types—it seems important to make a direct assessment of the extent to which subscores based on the three LSAT item types exhibit "differential test-retest stability." This might be done, for example, by using data for a sample of LSAT repeaters with differing time-intervals between test administrations. Questions regarding differential test-retest stability for logical reasoning and analytical reasoning subscores appear to be equally pertinent for the GRE context.

---

[16] Stricker (1992) examined mean discrepancies between pairs of item-type subscores on the current version of the LSAT, by gender, age, ethnicity, English-language background, and several other variables. Mean discrepancies between z-scaled mean on analytical reasoning, on the one hand, and reading comprehension and logical reasoning on the other hand, were relatively marked for older examinees (27 years of age or older)--who performed less well on the analytical reasoning subscore, and Asian Americans, and examinees reporting that they were not fluent in English--whose analytical reasoning performance was substantially better than their performance on either logical reasoning or reading comprehension. Differences were not analyzed by undergraduate major field.

[17] See Pitcher (1976b) for a conceptually similar study, concerned in part with whether differential weighting of the then extant section scores (Reading Comprehension, Reading Recall, Data Interpretation, and Principles and Cases) would improve prediction, and whether it might be useful to report separate section scores. Results for five schools indicated that the standard regression weight for Reading (combined RC and Recall) was larger than the weight for either Data Interpretation or Principles and Cases (an item type designed to measure the ability to reason logically, using items with legalistic content) for both men and women, but especially so for women.

It is of incidental interest to note that findings of this study tend to confirm and extend findings of GRE studies (such as those reviewed in Section 2), suggesting that logical reasoning and analytical reasoning items measure psychometrically distinguishable aspects of reasoning ability. The present findings and previous GRE findings also attest in common to an underlying affinity between logical reasoning and reading comprehension item types—both requiring the analysis of extended discourse or general reasoning; also to distinctions between aspects of general reasoning, on the one hand, and more narrowly constrained, formal-deductive aspects of reasoning that appear to be assessed by analytical reasoning items.

Study findings indicating that LSAT logical reasoning, reading comprehension, and analytical reasoning items and their GRE counterparts have a common factor structure, are important for both testing programs, because by inference these findings indicate that future research involving these item types in the LSAT context can validly draw on the findings of related GRE research in formulating working hypotheses and evaluating LSAT research outcomes. It follows that as LSAT research findings involving these item types accrue, the LSAT findings in turn can usefully inform research in the GRE context.

Regarding the observed affinity between logical reasoning and reading comprehension, it is pertinent to note that the version of the logical reasoning item type considered in this study involves heavy reading comprehension requirements. Accordingly, underlying differences between logical reasoning and reading comprehension, if present, may be confounded by differences in reading comprehension.

To the extent that it is possible to measure "logical reasoning" using item types with limited reading demands, progress may be made in clarifying distinctions between "logical reasoning" and "reading comprehension," in both LSAT and GRE testing contexts—for example, through cooperative research projects involving experimental item types, along lines suggested by Emmerich, Enright, Rock, and Tucker (1991).

In reaching decisions regarding score definition and score reporting, the results of this study involving data from both tests suggest that both testing programs are in a position to benefit from research projects capitalizing on the common structure that appears to underly performance on the types of items that are common to both the Law School Admission Test and the GRE General Test.

Finally, viewed from a purely methodological perspective, based on the orderly, interpretable outcomes of this study, the apparently novel procedures used in pooling data across diverse test forms for the purpose of generating correlations for analysis appear to warrant further research. As indicated earlier, these procedures involve an assumption of statistical equivalence for parcels of items of the same type and position in successive test forms. In addition, it was assumed that pooling z-scores across samples would have no appreciable effect on the resulting factor analytic structure.

Further research is needed to evaluate these assumptions, to assess the "statistical equivalence" of the parcels involved in this study, and to explore the utility of alternative approaches to pooling data across different test forms for the purpose of assessing internal construct validity and sensitivity of results to pooling across samples. The present LSAT/GRE data-base might be used to explore, for example, questions regarding effects of various degrees of departure from parallelism in item-type parcels on factor outcomes.

REFERENCES

Camilli, G., Wang, M-m., and Fesq, J. (1992). The effects of dimensionality on true score tables for the Law School Admission Test (LSAT Statistical Report 92-01). Newtown, PA: Law School Admission Services.

Carlson, A. B. (1976b). Factor analysis and validity study of the Law School Admission Test battery. Report #LSAC-70-3. In Law School Admission Council (1976b). Reports of LSAC Sponsored Research:Volume II, 1970-74. Princeton, NJ: Law School Admission Council.

Dorans, N. J., and Lawrence, I. M. (1987). The internal construct validity of the SAT (ETS RR-87-35). Princeton, NJ: Educational Testing Service.

Educational Testing Service (1990). GRE Information Bulletin, 1989-90. Princeton, NJ: Educational Testing Service.

Educational Testing Service (1992). Guide to the use of the Graduate Record Examinations Program, 1990-91. Princeton, NJ: Educational Testing Service.

Emmerich, W., Enright, M. K., Rock, D. A., and Tucker, C. (1991). The development, investigation, and evaluation of new item types for the GRE Analytical Measure (GRE Board Professional Report No. 87-09P and ETS Research Report 91-16). Princeton, NJ: Educational Testing Service.

Harman, H. H. (1976). Modern factor analysis. Chicago: University of Chicago Press.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

Kingston, N. M., and Dorans, N. J. (1982). The effect of the position of an item within a test on item responding behavior: An analysis based on item response theory. Gre Board Professional Report No. 79-12bP, ETS RR-82-22. Princeton, NJ: Educational Testing Service.

Law School Admission Test Papers, 1945-1973. Princeton, NJ: Educational Testing Service Archives.

Law School Admission Council (1976a). Reports of LSAC Sponsored Research: Volume I, 1949-1969. Princeton, NJ: Law School Admission Council.

Law School Admission Council (1976b). Reports of LSAC Sponsored Research: Volume II, 1970-74. Princeton, NJ: Law School Admission Council.

Law School Admission Council (1977). Reports of LSAC Sponsored Research: Volume III, 1975-77. Princeton, NJ: Law School Admission Council.

Law School Admission Council (1982). Law School Admission Test: Sources, Contents, and Uses. Newtown, PA: Law School Admission Services.

Law School Admission Council (1984). Reports of LSAC Sponsored Research: Vol. IV, 1978-83. Newtown, PA: Law School Admission Services.

Law School Admission Services (1990). Law Services Information Book, 1990-91. Newtown, PA: Law School Admission Services.

Law School Admission Services (1991). Law Services Information Book, 1990-91. Newtown, PA: Law School Admission Services.

McPeek, W. M., Pitcher, B., & Carlson, A. B. (1976b). The predictive effectiveness of several experimental item types and the operational item types in the Law School Admission Test. Report #LSAC-74-1. In Law School Admission Council, Reports of LSAC Sponsored Research: Volume II, 1970-74. Princeton, NJ: Law School Admission Council.

Pitcher, B. (1977a). LSAT part-score validity study. Report #LSAC-74-2. In Law School Admission Council (1977), Reports of LSAC Sponsored Research: Volume II, 1970-74. Princeton, NJ: Law School Admission Council.

Pitcher, B. (1977b). Subgroups validity study. Report #LSAC-76-6. In Law School Admission Council, Reports of LSAC Sponsored Research: Volume III, 1975-77. Princeton, NJ: Law School Admission Council.

Powers, D. E. (1988). A factor analytic investigation of seven experimental item types (GRE Board Professional Report 77-1P). Princeton, NJ: Educational Testing Service.

Powers, D. E., Swinton, S.S., and Carlson, A. B. (1977). A factor analytical study of the GRE Aptitude Test (GRE Board Professional Report 75-11P). Princeton, NJ: Educational Testing Service.

Powers, D. E., and Swinton, S. S. (1981). Extending the measurement of graduate admissions abilities beyond the verbal and quantitative domains. Applied Psychological Measurement, 5, 141-158.

Rock, D., Bennett, R. E., & Jirele, T. (1986). The internal construct validity of the GRE General Test across handicapped and nonhandicapped populations (Research Report No. 86-7). Princeton, NJ: Educational Testing Service.

Rock, D. A., Werts, C., and Grandy, J. (1982). Construct validity of the GRE Aptitude Test–An empirical confirmatory study (GRE 78-1P & ETS RR-81-57). Princeton, NJ: Educational Testing Service.

Rulon, P. J. (1939). A simplified procedure for determining the reliability of a test by split-halves, Harvard Educational Review, 9, 99-103.

Schaeffer, G. A., & Kingston, N. M. (1988). *Strength of the analytical factor of the GRE test in several subgroups: A full-information factor analysis approach* (GRE Board Professional Report No. 86-7P; also ETS Research Report 88-5). Princeton, NJ: Educational Testing Service.

Stricker, L. J., and Rock, D. A. (1985). *Factor structure of the GRE General Test for older examinees: Implications for construct validity* (GRE Board Research Report GREB No. 83-10R & ETS Research Report 85-9). Princeton, NJ: Educational Testing Service.

Stricker, L. J., & Rock, D. A. (1987). Factor structure of the GRE General Test in young and middle adulthood. *Developmental Psychology*, *23*, 526-536.

Swinton, S. S., and Powers, D. E. (1980). *A factor analytic study of the restructured GRE Aptitude Test* (GRE Board Professional Report 77-6P). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1988), A study of the long-term stability of GRE General Test Scores, *Research in Higher Education*, 29, 3-40.

Wilson, K. M. (1986). *The relationship of scores based on GRE General Test item types to undergraduate grades: An exploratory study for selected subgroups* (GRE Board Professional Report GREB No. 83-19P and ETS RR-86-37). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1985). *The relationship of GRE General Test item-type part scores to undergraduate grades* (GRE Board Professional Report GREB No. 81-22P & ETS Research Report 84-38). Princeton, NJ: Educational Testing Service.

Wightman, L. F., and Muller, D. G. (1990a). *Comparison of LSAT performance among selected subgroups* (Law School Admission Council Statistical Report 90-01). Newtown, PA: Law School Admission Services.

Wightman, L. F., and Muller, D. G. (1990b). *An analysis of differential validity and Differential Prediction for Black, Mexican American, Hispanic, and White law school students* (Law School Admission Council Research Report 90-03). Newtown, PA: Law School Admission Services.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (1/2003)